

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 14:46:42 ; Search time 22 Seconds
(Without alignments)
458,125 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 1325
Sequence: 1 MRRLVLLGLAGSPPLD.....DSTFGSLVSDMHSSPVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1325	100.0	243	1	COT5_HUMAN
2	433.5	32.7	247	1	APM1_MOUSE
3	424	32.0	244	1	APM1_HUMAN
4	418.5	31.6	680	1	CA1A_HUMAN
5	417	31.5	674	1	CA1A_CHICK
6	412	31.1	246	1	C10C_MOUSE
7	411.5	31.0	674	1	CA1A_BOVIN
8	410.5	31.0	289	1	COT7_HUMAN
9	408	30.8	680	1	CA1A_MOUSE
10	403	30.4	285	1	COT2_HUMAN
11	397	30.0	245	1	C10C_HUMAN
12	374	28.2	744	1	CA18_RABIT
13	373	28.2	419	1	COLE_MOUSE
14	368	27.8	744	1	CA18_HUMAN
15	366.5	27.7	743	1	CA18_MOUSE
16	363	27.4	635	1	CA28_HUMAN
17	362	27.3	251	1	C10B_HUMAN
18	360.5	27.2	253	1	C10B_RAT
19	350	26.4	253	1	C10B_MOUSE
20	314	23.7	245	1	C10A_HUMAN
21	314	23.7	245	1	C10A_MOUSE
22	298.5	22.5	255	1	GLIC_MOUSE
23	285.5	21.5	258	1	C10E_HUMAN
24	282	21.3	258	1	C10E_MOUSE
25	280.5	21.2	215	1	HP25_TAMSI
26	277.5	20.9	215	1	HP27_TAMSI
27	264.5	20.0	246	1	COT3_HUMAN
28	261.5	19.7	196	1	HP20_TAMSI
29	230.5	17.4	1049	1	CA13_BOVIN
30	229.5	17.3	281	1	COT1_HUMAN
31	228	17.2	684	1	CA39_HUMAN
32	224	16.9	636	1	CA13_RAT
33	223	16.8	1464	1	CA13_MOUSE

34	222.5	16.8	1758	1	CA24_CAEEL	P17140 caenorhabdit
35	221.5	16.7	170	1	CA28_MOUSE	P25318 mus musculus
36	221	16.7	1019	1	CA16_CHICK	P20785 gallus galli
37	221	16.7	1262	1	CA13_CHICK	P12105 gallus galli
38	220	16.6	248	1	PSPA_CANFA	P06908 canis familiaris
39	220	16.6	360	1	CCD2_CAEEL	P35799 caenorhabdit
40	219.5	16.6	247	1	PSPA_RABIT	P12842 oryctolagus
41	219.5	16.6	1466	1	CA13_HUMAN	P02461 homo sapien
42	219	16.5	671	1	CA11_RAT	P02454 rattus norv
43	218.5	16.5	1453	1	CA11_MOUSE	P11087 mus musculus
44	217.5	16.4	1516	1	CA1H_HUMAN	P39060 homo sapien
45	217	16.4	747	1	CA12_BOVIN	P02459 bos taurus

ALIGNMENTS

RESULT 1	ID	COT5_HUMAN	STANDARD:	PRT:	243 AA.
AC	09BXJ0: 09BXJ4:				
DT	15-JUN-2002 (Rel. 41, Created)				
DT	15-JUN-2002 (Rel. 41, Last sequence update)				
DT	15-JUN-2002 (Rel. 41, Last annotation update)				
DE	Complement-c1q tumor necrosis factor-related protein 5 precursor.				
GN	C1QTNF5 OR CTRP5.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	(1)				
RP	SEQUENCE FROM N.A.				
RA	Sheppard P.O., Humes J.M.;				
RT	"Homo sapiens complement-c1q tumor necrosis factor-related protein.";				
RL	Submitted (Dec-2000) to the EMBL/Genbank/DBJ databases.				
RP	(2)				
RN	SEQUENCE OF 25-243 FROM N.A.				
RC	Tissue-uterus;				
RA	Ottenswelder B., Obermayer B., Mewes H.-W., Gassenhuber J.,				
RA	Miemann S.;				
RL	Submitted (AUG-1999) to the EMBL/Genbank/DBJ databases.				
CC	-1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.				
CC	-1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).				
CC	-----				
DR	EMBL: AF329841: AAK17965.1: -				
DR	EMBL: AL110261: CAB3702.1: -				
DR	GeneW: HGNC:14344: C1QTNF5.				
DR	InterPro: IPR001073: C1q.				
DR	InterPro: IPR000087: Collagen.				
DR	Pfam: PF00386: C1q: 1.				
DR	Pfam: PF01391: Collagen: 1.				
DR	PRINTS: PRO0007: COMPLEMENTC1Q.				
DR	SMART: SM00110: C1Q: 1.				
DR	PROSITE: PS01113: C1Q: FALSE_NEG.				
KW	Collagen; Signal.				
FT	SIGNAL	1	15	POTENTIAL.	
FT	CHAIN	16	243	COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-	
FT	DOMAIN	30	95	RELATED PROTEIN 5.	
FT	DOMAIN	97	243	COLLAGEN-LIKE.	
SO	SEQUENCE	243 AA;	25298 MW;	70CDA65CDA7EB784 CRC64;	

Query Match: 100.0%; Score 1325; DB 1; Length 243;
Best Local Similarity: 100.0%; Pred. No. 1.7e-85;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

AC Q13896-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Adiponectin precursor (30 kDa adipocyte complement-related protein)
DE (ACRP30) (Adipose most abundant gene transcript 1) (ap4-1) (gelatin-
DE binding protein)
DE
GN APM1 OR ACRP30 OR GBP28.
OS Homo sapiens (Human).
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Adipose tissue;
RX MEDLINE=96224171; Pubmed=8619847;
RA Maeda K., Okubo K., Shimomura I., Funahashi T., Matsuzawa Y.,
RA Matsubara K.;
RT "cDNA cloning and expression of a novel adipose specific collagen-like
RT factor, ap41 (Adipose most abundant Gene transcript 1).";
RL Biochem. Biophys. Res. Commun. 221:286-289(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99136984; Pubmed=10095105;
RA Saito K., Tobe T., Minoshima S., Asakawa S., Sumiya J., Yoda M.,
RA Nakano Y., Shimizu N., Tomita M.;
RT "Organization of the gene for gelatin-binding protein (GBP28).";
RL Gene 229:67-73(1999).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=99333693; Pubmed=10403784;
RA Schaeffler A., Orso E., Paltzsch K.D., Buechler C., Drobnik W.,
RA Fuerst A., Schoelmerich J., Schmitz G.;

CC This SWISS-PROT entry is copyright. It is produced through a collab
CC

-1- SIMILARITY: CONTAINS 1 CID CONTAIN:

This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch)
 CC -----
 DR EMBL: DA5371; BAA08227.1; -
 DR EMBL: AB012165; BAA86716.1; -
 DR EMBL: AB012164; BAA86716.1; JOINED.
 DR EMBL: AJ131460; CAB52413.1; -
 DR EMBL: AJ131461; CAB52413.1; JOINED.
 DR MIM: 605441; -
 DR InterPro: IPR001073; C1Q.
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF00386; C1Q; 1.
 DR Pfam: PF01391; Collagen; 1.
 DR PRINTS: PRO0007; COMPLEMENTC1Q.
 DR SMART: SM00110; C1Q; 1.
 DR PROSITE: PS01113; C1Q; 1.
 KW Hormone; Collagen; Signal; Repeat; Hydroxylation; Plasma;
 KW Polymorphism; Disease mutation; Obesity; Diabetes mellitus.
 FT SIGNAL 1 14
 FT CHAIN 15 244
 FT DOMAIN 42 107
 FT DISULFID 36 244
 FT MOD_RES 44 44 INTERCHAIN (BY SIMILARITY).
 FT MOD_RES 47 47 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 53 53 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 62 62 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 71 71 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 76 76 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 86 86 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 95 95 HYDROXYLATION (BY SIMILARITY).
 FT MOD_RES 104 104 HYDROXYLATION (BY SIMILARITY).
 FT VARIANT 84 84 G -> R.
 FT VARIANT 112 112 R -> C (IN ADIPONECTIN DEFICIENT).
 FT VARIANT 117 117 V -> M.
 FT VARIANT 164 164 I -> T.
 FT VARIANT 221 221 R -> S.
 FT VARIANT 241 241 H -> P.
 FT SEQUENCE 244 AA; 26414 MW; 64806C61204B1018 CRC64;
 Query Match 32.0%; Score 424; DB 1; Length 244;
 Best Local Similarity 40.5%; Pred. No. 5.3e-23;
 Matches 106; Conservative 26; Mismatches 72; Indels 58; Gaps 10;
 OY 6 VLLLLGLAAGSPRLDDNKIPSLCPH-----PG-----LGTGPHHSQ 44
 DB 7 VLLLLAL-----PCHDDETTGGVLLPLPKGACGTGMMAGIPGHPHN 50
 OY 45 GLPGRDGDRDGAAGAPGEGGEGRGLPGRGCD-----PGRGEAGRAPPTPA 95
 DB 51 GAPRGDGD-----GTPGEGEKGDPCLLIGPKGIDIGETGVPAGACPGPGFIOGRKEP 104
 OY 96 GECVVPASAFSAKRSRPPSPDAPLPFRVLVNEQGHDAVTKFTQVPGVYFPAV 155
 DB 105 GEGAVVYSASFV-GLETVYVIP-NMPLRTKIFENQONHNDGSGTKFHCNIPGLYFPAV 162
 OY 156 HATVYRASLOPDLVKNGES-IASFQFEGGWPKPASLSGAMVRLPEPDQVWVQV-CVGD 213
 DB 163 HITVYMKVKSLEFKDKAMLPFTDQYQENNVDA--SGSVLLHLEVDQWLVQYGE 220
 OY 214 YIGIVASIKTDSFGSLVSD 235
 DB 221 RNLGYADNDNDSTFGFLYHD 242

RESULT 4
 C1A1_HUMAN STANDARD: PRT: 680 AA.
 AC 003692; 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Collagen alpha 1(X) chain precursor.
 GN COL10A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OC NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-92109659; PubMed-1764025;
 RA Thomas J.T., Cresswell C.J., Rash B., Nicolai H., Jones T.,
 RA Solomon E., Grant M.E., Boot-Handford R.P.;
 RT "The human collagen X gene. Complete primary translated sequence and
 RT chromosomal localization.";
 RL Biochem. J. 280:617-623(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-93012005; PubMed-1397333;
 RA Reichenberger E., Beier F., Luvalle P., Olsen B.R., von der Mark K.,
 RA Bertling W.M.;
 RT "Genomic organization and full-length cDNA sequence of human collagen
 RT X.";
 RL FEBS Lett. 311:305-310(1992).
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Beier F., Lamm M.B., von der Mark K.;
 RL Submitted (JUN-1996) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Williams S.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 52-680 FROM N.A.
 RX MEDLINE-92267014; PubMed-1587271;
 RA Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;
 RT "Cloning of the human and mouse type X collagen genes and mapping of
 RT the mouse type X collagen gene to chromosome 10.";
 RL Eur. J. Biochem. 206:217-224(1992).
 RN [6]
 RP SEQUENCE OF 561-666 FROM N.A.
 RX MEDLINE-91243838; PubMed-2037056;
 RA Apte S., Mattei M.G., Olsen B.R.;
 RT "Cloning of human alpha 1(X) collagen DNA and localization of the
 RT COL10A1 gene to the q21-q22 region of human chromosome 6.";
 RL FEBS Lett. 282:393-396(1991).
 RN [7]
 RP SEQUENCE OF 547-655 FROM N.A.
 RX MEDLINE-92077285; PubMed-1743401;
 RA Reichenberger E., Aigner T., von der Mark K., Stoeb H., Bertling W.;
 RT "In situ hybridization studies on the expression of type X collagen
 RT in fetal human cartilage.";
 RL dev. Biol. 148:562-572(1991).
 RN [8]
 RP REVIEW ON VARIANTS.
 RX MEDLINE-97255959; PubMed-9101290;
 RA Kuivalainen H., Tromp G., Prockop D.J.;
 RT "Mutations in fibrillar collagens (types I, II, and XI), fibril-
 RT associated collagen (type IX), and network-forming collagen (type X)
 RT cause a spectrum of diseases of bone, cartilage, and blood vessels.";
 RL Hum. Mutat. 9:300-315(1997).
 RN [9]
 RP VARIANTS SMCD ASP-598 AND PRO-614.
 RX MEDLINE-94136476; PubMed-8304336;
 RA Wallis G.A., Rash B., Sweetman W.A., Thomas J.T., Super M., Evans G.,
 RA Grant M.E., Boot-Handford R.P.;

DB 521 MPECFIKAGRPRLSGTPLVSNAGVGMVSAFTVILSKAY--PAITPTIPDKILLYNR 578
 QY 133 QGHDAVATGKFTCQVPGVYFAVHATYVRASLQFDLVKNGESIA-SFFQFGGPKPASTL 191
 DB 579 QGHDPKRCITTCQIPGTYFSYHVGCHTVMVGLNGTIPVMTYDEYTKGYLDA-- 636
 QY 192 SGGANVRLPEPDQVWVGVDYIGIVASIKTDSFGFLV 232
 DB 637 SGSAVIDLTENDQWVLQPLNAESNGLYSSEYVSSFGFLV 677

RESULT 5
 CAIA_CHICK STANDARD: PRT: 674 AA.
 ID CAIA_CHICK P08125;
 AC 01-AUG-1988 (Rel. 08, Created)
 DT 01-NOV-1991 (Rel. 20, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Collagen alpha 1(X) chain precursor.
 GN COL10A1.
 OS Gallus gallus (Chicken).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
 OC Gallus.
 OX NCBI_TaxID=9031;
 RN RP SEQUENCE OF 48-674 FROM N.A., AND SEQUENCE OF 103-117 AND 453-466.
 RA MEDLINE=86168227; PubMed=3082876;
 RA Niinomiya Y., Gordon M., van der Rest M., Schmid T., Linsenmayer T.,
 RA Olsen B.R.,
 RT "The developmentally regulated type X collagen gene contains a long
 RT open reading frame without introns."
 RT J. Biol. Chem. 261:5041-5050(1986).
 RN RP SEQUENCE OF 1-75 FROM N.A.
 RA MEDLINE=89054019; PubMed=2461368;
 RA Luvallie P., Niinomiya Y., Rosenblum N.D., Olsen B.R.,
 RT "The type X collagen gene. Intron sequences split the 5'-untranslated
 RT region and separate the coding regions for the non-collagenous amino-
 RT terminal and triple-helical domains."
 RT J. Biol. Chem. 263:18378-18385(1988).
 RN RP [3]
 RP REVISIONS TO C-TERMINUS.
 RX MEDLINE=89380199; PubMed=2476437;
 RA Yamaguchi N., Benya P.D., van der Rest M., Niinomiya Y.,
 RT "The cloning and sequencing of alpha 1(VIII) collagen cDNAs.
 RT demonstrate that type VIII collagen is a short chain collagen and
 RT contains triple-helical and carboxyl-terminal non-triple-helical
 RT domains similar to those of type X collagen."
 RT J. Biol. Chem. 264:16022-16029(1989).
 CC -1- FUNCTION: TYPE X COLLAGEN IS A PRODUCT OF HYPERTROPHIC
 CC CHONDROCYTES AND HAS BEEN LOCALIZED TO PRESUMPTIVE
 CC MINERALIZATION ZONES OF HYALINE CARTILAGE.
 CC -1- SUBUNIT: HOMOTRIMER.
 CC -1- PMW: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -1- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
 CC -1- SIMILARITY: STRONG, TO ALPHA 1 AND 2 TYPE VIII COLLAGENS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC EMBL: M13496; AAA48736.1; ALT_SEQ.
 DR EMBL: J04194; AAA48634.1; -.
 DR PIR: A31896; A31896.
 DR InterPro: IPR001073; C1q.
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF00386; C1q; 1.

DR Pfam: PF01391; Collagen; 8.
 DR PRINTS: PRO0007; COMPLEMENTC1Q.
 DR SMART: SM00110; C1Q; 1.
 DR PROSITE: PS01113; C1Q; 1.
 KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
 KW Cartilage; Collagen; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 674 COLLAGEN ALPHA 1(X) CHAIN.
 FT DOMAIN 19 52 NONHELICAL REGION (NC2).
 FT 53 512 TRIPLE-HELICAL REGION.
 FT DOMAIN 513 674 NONHELICAL REGION (NC1).
 FT 539 674 C1Q.
 FT DOMAIN 453 453 HYDROXYLATION.
 FT MOD_RES 456 456 HYDROXYLATION.
 FT MOD_RES 456 456
 SQ SEQUENCE 674 AA; 66434 MW; EAB48B1EF174B145 CRC64;

Query Match 31.5%; Score 417; DB 1; Length 674;
 Best Local Similarity 36.0%; Pred. No. 4.4e-22;
 Matches 96; Conservative 34; Mismatches 84; Indels 53; Gaps 6;

QY 14 AGSPPLDNNKIPSLCPGHPGLPTGHHGSGQLGRDGRDGAAPGAPGEGGAPGL 73
 DB 408 AGHPGLGPRVPGQGVKVPNGINERPGPSGIRGVGPIGPPGMPGAPGAKGAPGL 467
 QY 74 PGPR-----GDGPRGEA-----GPAPTGPAGECSVP----- 101
 DB 468 PGPAIVTKGLRGPMPGLGPRGKSGDEBQLPGPPPPPPPGPGOSTIPEGYVKGESNELS 527
 QY 102 -----PRSAFSAKRSERVPSPDAPLPDRVVLVNEQHYDAVTKFTC 145
 DB 528 GMSMKKAGANALTMGVSAFTVILSKAY--PGATVIRKDKILYNGQHYDPTGFTC 565
 QY 146 QVPGVYFAVHATYVRASLQFDLVKNGESIA-SFFQFGGPKPASTL 204
 DB 586 RLPGLYFSYHVGCHTVMVGLNGTIPVMTYDEYTKGYLDA--SGSAVIDLTENDQ 643
 QY 205 VVWGVGVGDYIGIVASIKTDSFGFL 231
 DB 644 VWLQPLNPSNGLYSSEYVSSFGFL 670

RESULT 6
 C1QC_MOUSE STANDARD: PRT: 246 AA.
 ID C1QC_MOUSE
 AC 002105;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Complement C1q subcomponent, C chain precursor.
 GN C1QG OR C1QC.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN RP [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-ICR; TISSUE-Macrophage;
 RC MEDLINE=93011118; PubMed=1396691;
 RA Petry F., Reid K.B.M., Loos M.,
 RT "Isolation, sequence analysis and characterization of cDNA clones
 RT coding for the C chain of mouse C1q. Sequence similarity of
 RT complement subcomponent C1q, collagen type VIII and type X and
 RT precerebellin."
 RT Eur. J. Biochem. 209:129-134(1992).
 RN RP [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c; TISSUE-Liver;
 RC MEDLINE=96186528; PubMed=8606057;
 RA Petry F., Mcclive P.J., Botta M., Morley B.J., Morahan G., Loos M.,
 RT "The mouse C1q genes are clustered on chromosome 4 and show
 RT conservation of gene organization."
 RT Immunogenetics 43:370-376(1996).
 CC -1- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYME C1R AND C1S TO YIELD

Matches 96; Conservative 34; Mismatches 84; Indels 61; Gaps 6;

QY 15 GSPPLDNDKIPSLCPGHPGLPTGPHGSGGLRGDRDGDGAPGKGGGRGAP 74
 Db 401 GNDGLGPKRDPGLAGSPGLPGVPGAGACGPHNGGAGPRVPLGTRGIGGPG 460
 QY 75 -----GPRGDPGRCGA-----GAPPTGPAEGCS 99
 Db 461 GPGSGKDVGTGPPGPPAGIAGKGLNGPLGPPGPPGAGPPGLPGPPGPPGQVA 520
 QY 100 VP-----PSAFSAKRSRSPPSADPLPDDVLYVMEGCHYA 138
 Db 521 LPEDFVKAGORPVANSAGVGTGMPVSAFYVILSKAT--PALGTPDPDLINKQDHYD 578
 QY 139 VGGKFCQVPGVYFAVHATVYRASLQDFLVKNGESIA-SFQFFGSGMPKPSLSGAMV 197
 Db 579 RGIFGCKIPGIYFYSYHIVKGTAMVGLKNGFVMTYDEYIKGYLDAQ--SSANVI 636
 QY 198 RLEPDQVWVGVGVGIGIVASIKTDSFFSGFLV 232
 Db 637 DLTENDQVWLQPLMNGSNGLYSPRYVHSSFGFLV 671

RESULT 8

CQT7_HUMAN STANDARD; PRT; 289 AA.

AC 09BXJ2;
 DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DE 15-JUN-2002 (Rel. 41, Last annotation update)
 GN Complement-clq tumor necrosis factor-related protein 7 precursor.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 NCBI_TaxID=9606;
 RN [1]
 RA Piddington C.S., Sheppard P.O., Bishop P., Lasser G.W.;
 RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RA "SEQUENCE FROM N.A."
 RC TISSUE=Testis;
 RA Strausberg R.;
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.1sb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL: AF398839; AKR17963.1;
 DR EMBL: BC022187; AAR22187.1;
 DR Genbank: HGNC:14342; C1QTNF7.
 DR InterPro: IPR001073; C1Q.
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF00386; C1q; 1.
 DR Pfam: PF01391; Collagen; 2.
 DR PRINTS: PRO0007; COMPLEMENTC1Q.
 DR SMART: SM00110; C1Q; 1.
 DR PROSITE: PS01113; C1Q; 1.
 DR Collagen; Signal.
 KW COLLAGEN; Signal.
 FT CHAIN 1 16
 FT SIGNAL 17 289
 FT DOMAIN 38 139
 FT DOMAIN 141 276
 FT C1Q.
 FT POTENTIAL.
 FT COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
 FT RELATED PROTEIN 7.
 FT COLLAGEN-LIKE.
 FT C1Q.

SEQUENCE 289 AA; 30683 MW; A61609FF68D26946 CRC64;

Query Match 31.0%; Score 410.5; DB 1; Length 289;

Best Local Similarity 40.0%; Pred. No. 5,4e-22;

Matches 98; Conservative 26; Mismatches 84; Indels 37; Gaps 7;

QY 24 IPSL--CPGHPGLPTGPHGSGGLRGDRDGDGAPGKGGGRGAP-----GAPGKG 66
 Db 36 IPGLGPPGPPGANSPPGPHGRTIGLGRDGRGRGKGEKGTAGLRCKTGPLGLAGEKG 95
 QY 67 EGGRGGLDPPRGDPPRGEGAGPAGTPGAGE-----C-SVPSAFSAKRS 111
 Db 96 DQGETGKKQPIGPPGEGKEGVGIPGPPGKCDRGEGDGLPGVGRCSIVLAKSAFSGIT 155
 QY 112 ESRVPPPSDAPLPFRVLVNEGHDVATGKFTQVPGVYFAVHATVYRASLQDFLVKN 171
 Db 156 TSY--PEERLPLTFKVLNFGNEHNPATGKFAFPGIYFYSYDITLANKHLAIGLVHN 213
 QY 172 GESIASFFQFGGMPKPSLSGAMVRLPDDQVWVGVGVGDIYIVASIK-TDSTFSGF 230
 Db 214 GQYRKITPDANTGNHDA--SGSYIYILQPDVWLEIFFTDQGLFSDPGMADSLFSGF 271
 QY 231 LVYSD 235
 Db 272 LLYVD 276

RESULT 9

CA1A_MOUSE STANDARD; PRT; 680 AA.

ID CA1A_MOUSE

AC 005306;

DT 01-NOV-1995 (Rel. 32, Created)

DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 01-NOV-1995 (Rel. 32, Last annotation update)

DE Collagen alpha 1(X) chain precursor.

DE COL10A1.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Scurionath; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RN [1]

RA "The mouse collagen X gene: complete nucleotide sequence, exon

RT structure and expression pattern.";

RL Biochem. J. 289:247-253(1993).

RN [2]

RA "SEQUENCE FROM N.A."

RC STRAIN=BA1B/C;

RC MEDLINE=93143676; PubMed=8424763;

RA Elima K., Eerola I., Rosati R., Metsaranta M., Garofalo S., Perala M.,

de Crombrughe B., Vuorio E.;

"The mouse collagen X gene: complete nucleotide sequence, exon

RT structure and expression pattern.";

RL Biochem. J. 289:247-253(1993).

RN [3]

RA "SEQUENCE OF 51-680 FROM N.A."

RC STRAIN=BA1B/C;

RC MEDLINE=92267014; PubMed=1587271;

RA Apte S.S., Seldin M.F., Hayashi M., Olsen B.R.;

"Cloning of the human and mouse type X collagen genes and mapping of

RT the mouse type X collagen gene to chromosome 10.";

RL Eur. J. Biochem. 206:217-224(1992).

RN [4]

RA "SEQUENCE OF 385-627 FROM N.A."

RC STRAIN=C57BL/6;

RC MEDLINE=92182017; PubMed=1543751;

RA Elima K., Metsaranta M., Kallio J., Peralae M., Eerola I.,

Garofalo S., de Crombrughe B., Vuorio E.;

"Specific hybridization probes for mouse alpha 2(IX) and alpha 1(X)

RT

Db 131 KKEPGLPGSCSGSGHTKSAFSAVATKSYRRELPKIFDKILMNEGHYNASSGKPYVC 190
QY 147 VPGVYFAVAATYVYRASLODLYKNCESTAFQFQFGMKPKPSLSCGANVRLPEQVW 206
Db 191 VPGVYFTYDTILANKHLAELVHNCQYRIKRTDANTGNHVA--SGSTILALKGDEW 248
QY 207 VQGVGVGYIGI-YASIKTDSTFSGFLVYSD 235
Db 249 LQIFYSQNGLEFYDPYWTDSLFTGFLITAD 278

RESULT 11
ID C1QC_HUMAN STANDARD: PRT: 245 AA.
C1QC_HUMAN
21-JUN-1986 (Rel. 01, Created)
21-JUN-2002 (Rel. 01, Last sequence update)
15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement C1q subcomponent, C chain precursor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Monocytes;
RX MEDLINE=91174759; PubMed=1706597;
RA Seliger G.C., Blake D.J., Reid K.B.M.;
RT "Characterization and organization of the genes encoding the A-, B-
and C-chains of human complement subcomponent C1q. The complete
RT derived amino acid sequence of human C1q.";
RL Biochem. J. 274:481-490(1991).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Cerebellum;
RA Nishit T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Suzuki Y.,
Hata H., Nakagawa K., Mizuno S., Morinaga M., Kawamura M.,
Sugiyama T., Irie R., Otsuki T., Sato H., Nishikawa T., Sugiyama A.,
Kawakami B., Nagai K., Isogai T., Sugano S.;
RT "NEDO human cDNA sequencing project.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
[4]
RP SEQUENCE OF 29-122.
RX MEDLINE=80020137; PubMed=486087;
RA Reid K.B.M.;
RT "Complete amino acid sequences of the three collagen-like regions
RT present in subcomponent C1q of the first component of human
RT complement.";
RL Biochem. J. 179:367-371(1979).
[5]
RP REVIEW OF C1Q DEFICIENCY.
RX MEDLINE=98450587; PubMed=9777412;
RA Petry F.;
RT "Molecular basis of hereditary C1q deficiency.";
RL Immunobiology 199:286-294(1998).
-1- FUNCTION: C1Q ASSOCIATES WITH THE PROENZYMES C1R AND C1S TO YIELD
C1, THE FIRST COMPONENT OF THE SERUM COMPLEMENT SYSTEM. THE
COLLAGEN-LIKE REGIONS OF C1Q INTERACT WITH THE CA(2+)-DEPENDENT
C1R(2)C1S(2) PROENZYME COMPLEX, AND EFFICIENT ACTIVATION OF C1
TAKES PLACE ON INTERACTION OF THE GLOBULAR HEADS OF C1Q WITH THE
FC REGIONS OF IGG OR IGM ANTIBODY PRESENT IN IMMUNE COMPLEXES.
-1- SUBUNIT: C1 IS A CALCIUM-DEPENDENT TRIMOLECULAR COMPLEX OF C1Q, R
AND S IN THE MOLAR RATION OF 1:2:2. C1Q SUBCOMPONENT IS COMPOSED
OF NINE SUBUNITS, SIX OF WHICH ARE DISULFIDE-LINKED DIMERS OF THE
A AND B CHAINS, AND THREE OF WHICH ARE DISULFIDE-LINKED DIMERS OF
THE C CHAIN.
-1- PTM: O-LINKED GLYCANS CONSIST OF GLC-GAL DISACCHARIDES.

CC -1- DISEASE: DEFECTS IN C1Q ARE A CAUSE OF C1Q DEFICIENCY. IT IS A
CC RARE GENETIC DISORDER WHICH IS ASSOCIATED WITH RECURRENT
CC INFECTIONS AND A HIGH PREVALENCE OF LUPUS ERYTHEMATOSUS-LIKE
CC SYMPTOMS. IT IS CHARACTERIZED BY A LOSS OF ACTIVATION OF THE
CC COMPLEMENT CLASSICAL PATHWAY.
CC -1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AK057792; BAB71575.1;
CC EMBL: BC009016; AAH09016.1;
CC PIR: A03207; C1HUOC.
CC PIR: S14351; S14351.
CC Genew: HGNC:1245; C1QC.
CC MIM: 120575;
CC InterPro: IPR001073; C1q.
CC InterPro: IPR000087; Collagen.
CC Pfam: PF00386; C1q; 1.
CC PRINTS: PR00007; COMPLEMENTC1Q.
CC SMART: SM00110; C1Q; 1.
CC DR POSITE: PS01113; C1Q; 1.
CC KM Complement pathway; Plasma; Hydroxylation; Glycoprotein; Collagen;
CC Repeat; Signal; Disease mutation.
FT SIGNAL 1 28
FT CHAIN 29 245
FT DOMAIN 31 112
FT DOMAIN 113 245
FT DISULFID 32 32
FT MOD_RES 35 36
FT MOD_RES 39 39
FT MOD_RES 42 42
FT MOD_RES 45 45
FT MOD_RES 54 54
FT MOD_RES 57 57
FT MOD_RES 63 63
FT MOD_RES 66 66
FT MOD_RES 71 71
FT MOD_RES 75 75
FT CARBOHYD 75 75
FT MOD_RES 81 81
FT MOD_RES 84 84
FT CARBOHYD 84 84
FT MOD_RES 93 93
FT MOD_RES 96 96
FT MOD_RES 99 99
FT MOD_RES 105 105
FT MOD_RES 105 105
FT VARIANT 43 43
FT CONFLICT 14 14
FT CONFLICT 23 23
FT CONFLICT 57 57
FT CONFLICT 66 66
FT CONFLICT 72 72
FT CONFLICT 84 84
FT CONFLICT 87 87
FT CONFLICT 90 90
FT CONFLICT 215 215
SQ SEQUENCE 245 AA; 25774 MW; FA1J117EB7ABFC12 CRC64;
Query Match 30.08; Score 397; DB 1; Length 245;
Best Local Similarity 39.88; Pred. No. 4e-21;
Matches 100; Conservative 33; Mismatches 84; Indels 34; Gaps 10;
QY 1 MRPLVLLILGLAAGSPRLDDNKRPSLC---PCHPLPCTPGHHSQGLPGRNGRGGRG 57
:::|||||I| : : | |||||I| | |||||I|


```

CC EMBL: U17431; AAA69978.1; ALT_FRAME.
DR InterPro: IPR001073; C1q.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; C1q; 1.
DR Pfam: PF01391; Collagen; 3.
DR PRINTS: PR00007; COMPLEMENTC1Q.
DR SMART: SM00110; C1Q; 1.
DR PROSITE: PS01113; C1Q; 1.
KW Extracellular matrix; Repeat; Collagen; Signal.
FT SIGNAL 1 19
FT CHAIN 20 419
FT DOMAIN 20 57
FT DOMAIN 58 274
FT DOMAIN 275 419
FT DOMAIN 272 419
FT CARBOHYD 37 37
FT CARBOHYD 320 320
FT CARBOHYD 419 AA; 43634 MW; 570CDB9675FC0F39 CRC64;
SQ SEQUENCE 419 AA; 43634 MW; 570CDB9675FC0F39 CRC64;

Query Match 28.28; Score 373; DB 1; Length 419;
Best Local Similarity 37.58; Pred. No. 3.1e-19;
Matches 96; Conservative 23; Mismatches 99; Indels 38; Gaps 6;

QY 15 GSPPLDNNK--IPSLCPGHGGLPCTPGHHSQGLRGR--DGRDRDAPAPRGKGGGR 70
DB 157 GEPGLNKTSGISGREGPMGRLACTKGLKGGLGCLGCKEKGRRPRGCGMGLNGT 216
QY 71 PGLRPGRRGRRP-----RGAPRAGPTGAGCSV 100
DB 217 DGVGGEGRGPRGLGCKGDTGARGPRGPGCGMGLGCKGKLVKRRPRGRRGPGSSVE 276
QY 101 PPRSAFSAKSSSESVPPPSDAPLPEDRVLVNKGCHDAVTKETQGVYFAVHAATVY 160
DB 277 QIRSAFSGVGLPEPSFPPSLPVKDFKVFYNGEGHMDPTLNKEFVTVPGVYLFSGHITVR 336
QY 161 RASIQFQLVNKG--ESIASPFQFGCGWPKRPSLSGCAVNRLEPEQVNVQVCGYIGIYA 219
DB 337 NRPRAALVYNGVRKKTROSLTGQDIDQA--SNLALLHLTGQVLETL--LRDMNGXYS 393
QY 220 SIKTSTSGFLVYSD 235
DB 394 SSEDSTSGFLLYPD 409

RESULT 14
CAB_HUMAN STANDARD; PRT; 744 AA.
AC P27658; Q96D07;
DT 01-AUG-1992 (Rel. 23, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen alpha 1(VIII) chain precursor (Endothelial collagen).
GN COL8A1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN 1[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91231001; PubMed=2029894;
RA Muragaki Y., Matel M.-G., Yamaguchi N., Olsen B.R., Nishimura Y.;
RT "The complete primary structure of the human alpha 1 (VIII) chain and
RT assignment of its gene (COL8A1) to chromosome 3.";
RL Eur. J. Biochem. 197;615-622(1991).
RN 1[2]
RP SEQUENCE FROM N.A.
RC Straussberg R.;
RC Tissue=Lung;
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: MAJOR COMPONENT OF THE DESCENDING MEMBRANE (BASEMENT
CC MEMBRANE) OF CORNEAL ENDOTHELIAL CELLS.
CC -1- SUBUNIT: MAY FORM HOMOTRIMERS, OR HETEROTRIMERS IN ASSOCIATION

```

```

CC WITH ALPHA 2(VIII) TYPE COLLAGENS.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -1- MISCELLANEOUS: FOUR CONSECUTIVE GLY-PRO-PRO TRIPLETS ARE PRESENT
CC AT THE C-TERMINUS OF THE TRIPLE-HELICAL REGION. THESE MAY PROVIDE
CC THE HIGH THERMAL STABILITY OF THIS REGION.
CC -1- SIMILARITY: STRONG, TO ALPHA 2 TYPES VIII AND X COLLAGENS.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL: X57527; CAA4748.1; -.
DR EMBL: BC013581; AAH13581.1; -.
DR PIR: S15435; S15435.
DR Genbank: HGNC:2215; COL8A1.
DR MIM: 120251; -.
DR InterPro: IPR001073; C1q.
DR InterPro: IPR000087; Collagen.
DR Pfam: PF00386; C1q; 1.
DR Pfam: PF01391; Collagen; 8.
DR PRINTS: PR00007; COMPLEMENTC1Q.
DR Prodom: PD000007; Collagen; 1.
DR SMART: SM00110; C1Q; 1.
DR PROSITE: PS01113; C1Q; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Cell adhesion; Collagen; Signal.
FT SIGNAL 1 28
FT CHAIN 29 744
FT DOMAIN 29 117
FT DOMAIN 118 571
FT DOMAIN 572 744
FT DOMAIN 609 744
FT DOMAIN 626 744
FT CONFLICT 297 297
FT CONFLICT 297 297
FT CONFLICT 344 344
FT CONFLICT 382 382
FT CONFLICT 382 382
FT CONFLICT 388 388
FT CONFLICT 454 454
FT CONFLICT 464 464
FT CONFLICT 601 601
FT CONFLICT 631 631
SQ SEQUENCE 744 AA; 73364 MW; 2BC1B0955DE2C9A3 CRC64;

Query Match 27.88; Score 368; DB 1; Length 744;
Best Local Similarity 33.38; Pred. No. 1.2e-18;
Matches 90; Conservative 34; Mismatches 84; Indels 62; Gaps 6;

QY 24 IPSLC--PGHPRGAPRGNHSGGLRGRDGRDGRDAPAPRGKGGGRGLP----- 74
DB 475 VPELLRPKRGPRGIGLQGLDGRGRLPGIGRSPRIGPRGIPGKGGGLRGPGRGIGK 534
QY 75 -----GPRGDPGRGEGAPGPTGP-----AGECS 99
DB 535 PGVAGLHGRPGKRGALPGQGRGRLPGRPRGPRGPRGPRGPRGPRGPRGPRGPRG 594
QY 100 VPPRSASFSAKSSSRP-----PPSDAPLPEDRVLVNKGCHDAVTKETQ 144
DB 595 VKPRHAYGAKKCGNGRPAVEMPAFTAEIAPRPVAPVAFKFNKLYNGQNYNPOTGIPT 654
QY 145 COVPGYVFAVHAATVYRASLDQVLNKGESIA--SFPQFGWPKRPSLSGCAVNRLEPE 203
DB 655 CEVPGYVFAVHAATVYRASLDQVLNKGESIA--SFPQFGWPKRPSLSGCAVNRLEPE 203
QY 204 QVWVGVGVVDYIGIVASIKTSTSGFLVY 233
DB 713 RVFLQWPSQAGLVAGVGVHSSFSGLLY 742

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 14:49:37 ; Search time 26 Seconds
(without alignments)
274.991 Million cell updates/sec

Title: US-09-943-851A-42

Perfect score: 1325

Sequence: 1 MRPLVLLVLLGLAAGSPPLD.....DSTFSGELVYSDWHSVPFA 243

Scoring table: BLOSUM62

Gapop 10.0 , Capext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2.6/prodata/1/laa/5A.COMB.pep.*
- 2: /cgn2.6/prodata/1/laa/5B.COMB.pep.*
- 3: /cgn2.6/prodata/1/laa/6A.COMB.pep.*
- 4: /cgn2.6/prodata/1/laa/6B.COMB.pep.*
- 5: /cgn2.6/prodata/1/laa/PCTUS.COMB.pep.*
- 6: /cgn2.6/prodata/1/laa/backfilest1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed.
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1325	100.0	243	4	US-09-140-804-2
2	1325	100.0	243	4	US-09-336-536-3
3	1271	95.9	243	4	US-09-188-930-295
4	1258	94.9	243	4	US-09-336-536-10
5	1256	94.8	228	4	US-09-336-536-11
6	1200	90.6	228	4	US-09-336-536-7
7	675	50.9	128	4	US-09-336-536-14
8	650	49.1	138	4	US-09-336-536-7
9	433.5	32.7	247	2	US-08-463-911-2
10	428.5	32.3	247	4	US-09-140-804-8
11	428.5	32.3	247	4	US-09-118-408-3
12	428.5	32.3	247	4	US-09-506-855-3
13	424	32.0	244	2	US-08-463-911-7
14	424	32.0	244	4	US-09-140-804-3
15	424	32.0	244	4	US-09-336-536-20
16	424	32.0	244	4	US-09-530-423-1
17	417	31.5	221	4	US-09-530-423-2
18	412	31.1	246	2	US-08-463-911-4
19	399	30.1	294	4	US-09-188-930-294
20	347	26.2	60	4	US-09-336-536-6
21	324	24.5	60	4	US-09-336-536-13
22	316	23.8	423	1	US-08-383-744-2
23	316	23.8	423	2	US-08-999-336-2
24	316	23.8	423	5	PCT-US96-01427-2
25	280.5	21.2	215	4	US-09-140-804-5
26	280.5	21.2	215	4	US-09-140-804-5
27	277.5	20.9	185	2	US-08-463-911-3

28	276	20.8	236	4	US-09-140-804-6	Sequence 6, Appl1
29	266.5	20.1	198	4	US-09-188-930-138	Sequence 138, App
30	257.5	19.4	130	4	US-09-485-316A-13	Sequence 13, Appl
31	251.5	19.0	130	4	US-09-485-316A-12	Sequence 12, Appl
32	243.5	18.4	130	4	US-09-485-316A-11	Sequence 11, Appl
33	240	18.1	161	4	US-09-415-551-3	Sequence 3, Appl1
34	229.5	17.3	280	4	US-09-247-155-178	Sequence 178, App
35	229.5	17.3	281	4	US-09-118-408-2	Sequence 2, Appl1
36	229.5	17.3	281	4	US-09-506-855-2	Sequence 2, Appl1
37	228	17.2	684	1	US-08-555-669-12	Sequence 12, Appl
38	228	17.2	684	3	US-09-073-663-12	Sequence 12, Appl
39	224	16.9	228	4	US-09-219-849-38	Sequence 38, Appl
40	223	16.8	281	4	US-09-118-408-44	Sequence 44, Appl
41	223	16.8	281	4	US-09-506-855-44	Sequence 44, Appl
42	221	16.7	124	4	US-09-485-316A-9	Sequence 9, Appl1
43	219.5	16.6	1057	3	US-08-931-820-4	Sequence 4, Appl1
44	219.5	16.6	1078	4	US-08-963-825-21	Sequence 21, Appl
45	219.5	16.6	1078	4	US-09-500-811-21	Sequence 21, Appl

ALIGNMENTS

```
RESULT 1
US-09-140-804-2
: Sequence 2, Application US/09140804
: Patent No. 6197930
: GENERAL INFORMATION:
: APPLICANT: Sheppard, Paul O.
: TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
: FILE REFERENCE: 97-49
: CURRENT APPLICATION NUMBER: US/09/140, 804
: CURRENT FILING DATE: 1998-08-26
: EARLIER APPLICATION NUMBER: 60/056, 983
: EARLIER FILING DATE: 1997-08-26
: NUMBER OF SEQ ID NOS: 47
: SOFTWARE: FastSeq for Windows Version 3.0
: SEQ ID NO 2
: LENGTH: 243
: TYPE: PRT
: ORGANISM: Homo sapiens
: US-09-140-804-2

Query Match      100.0%; Score 1325; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 9, 5e-114;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 MRPLVLLVLLGLAAGSPPLDKNKIPSLCPGHPGLPTPGHHGSGGLPGRDGRDGPAG 60
DB      1 MRPLVLLVLLGLAAGSPPLDKNKIPSLCPGHPGLPTPGHHGSGGLPGRDGRDGPAG 60

QY      61 APGEKGGGRGRLPGPRGDGPRGEGAPGPTGPGAGCSVPPRSASFSAKSESIVPPSD 120
DB      61 APGEKGGGRGRLPGPRGDGPRGEGAPGPTGPGAGCSVPPRSASFSAKSESIVPPSD 120

QY      121 APLEFDVNLVNEGCHYAVYVGTQGVYFFVAVHTVVRASIQFDLVNKGESIASFFQ 180
DB      121 APLEFDVNLVNEGCHYAVYVGTQGVYFFVAVHTVVRASIQFDLVNKGESIASFFQ 180

QY      181 FFGGMPRPASISGCAWRLPEPDQVWQVGVGDVIGIYASIKTSTFSGFLVYSMDWSSP 240
DB      181 FFGGMPRPASISGCAWRLPEPDQVWQVGVGDVIGIYASIKTSTFSGFLVYSMDWSSP 240

QY      241 VFA 243
DB      241 VFA 243

RESULT 2
US-09-336-536-3
: Sequence 3, Application US/09336536
: Patent No. 6406884
```

GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 3
LENGTH: 243
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-3

Query Match 100.0%; Score 1325; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 9,5e-114;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHSGGLPGRDGRDGRDAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHSGGLPGRDGRDGRDAPG 60
QY 61 APGEKGGGRPGLPGRGDPGRGEGAPGACPTGACGCSVPPRSASFSAKRSSESVPPSPD 120
DB 61 APGEKGGGRPGLPGRGDPGRGEGAPGACPTGACGCSVPPRSASFSAKRSSESVPPSPD 120
QY 121 APLEPDRVLNDEGHDAVTGKFTCOVPGYVFAVHATVYRASLOPDLVKNGSIASFFQ 180
DB 121 APLEPDRVLNDEGHDAVTGKFTCOVPGYVFAVHATVYRASLOPDLVKNGSIASFFQ 180
QY 181 FFGWMPKPAASLSCGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFGFLVYSDMHSSP 240
DB 181 FFGWMPKPAASLSCGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFGFLVYSDMHSSP 240

QY 241 VFA 243
DB 241 VFA 243

RESULT 3
US-09-188-930-295
Sequence 295, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Streahan, Lotna
APPLICANT: Sleeman, Matthew
APPLICANT: Murison, Rene
TITLE OF INVENTION: Compositions Isolated from Skin Cells
FILE REFERENCE: 11000,1011C1
CURRENT APPLICATION NUMBER: US/09/188,930A
CURRENT FILING DATE: 1998-11-09
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 295
LENGTH: 243
TYPE: PRT
ORGANISM: Rat
US-09-188-930-295

Query Match 95.9%; Score 1271; DB 4; Length 243;
Best Local Similarity 95.1%; Pred. No. 8,2e-109;
Matches 231; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHSGGLPGRDGRDGRDAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHSGGLPGRDGRDGRDAPG 60
QY 61 APGEKGGGRPGLPGRGDPGRGEGAPGACPTGACGCSVPPRSASFSAKRSSESVPPSPD 120

DB 61 APGEKGGGRPGLPGRGDPGRGEGAPGACPTGACGCSVPPRSASFSAKRSSESVPPSPD 120
QY 121 APLEPDRVLNDEGHDAVTGKFTCOVPGYVFAVHATVYRASLOPDLVKNGSIASFFQ 180
DB 121 APLEPDRVLNDEGHDAVTGKFTCOVPGYVFAVHATVYRASLOPDLVKNGSIASFFQ 180
QY 181 FFGWMPKPAASLSCGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFGFLVYSDMHSSP 240
DB 181 FFGWMPKPAASLSCGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 4
US-09-336-536-10
Sequence 10, Application US/09336536
Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 10
LENGTH: 243
TYPE: PRT
ORGANISM: Mus musculus
US-09-336-536-10

Query Match 94.9%; Score 1258; DB 4; Length 243;
Best Local Similarity 93.8%; Pred. No. 1,3e-107;
Matches 228; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHSGGLPGRDGRDGRDAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHSGGLPGRDGRDGRDAPG 60
QY 61 APGEKGGGRPGLPGRGDPGRGEGAPGACPTGACGCSVPPRSASFSAKRSSESVPPSPD 120
DB 61 APGEKGGGRPGLPGRGDPGRGEGAPGACPTGACGCSVPPRSASFSAKRSSESVPPSPD 120
QY 121 APLEPDRVLNDEGHDAVTGKFTCOVPGYVFAVHATVYRASLOPDLVKNGSIASFFQ 180
DB 121 APLEPDRVLNDEGHDAVTGKFTCOVPGYVFAVHATVYRASLOPDLVKNGSIASFFQ 180
QY 181 FFGWMPKPAASLSCGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFGFLVYSDMHSSP 240
DB 181 FFGWMPKPAASLSCGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 5
US-09-336-536-4
Sequence 4, Application US/09336536
Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
APPLICANT: Bossone, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
CURRENT FILING DATE: 1999-06-18

NUMBER OF SEQ ID NOS: 75
 : SOFTWARE: Patentln Ver. 2.0
 : SEQ ID NO 4
 : LENGTH: 228
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-336-536-4

Query Match 94.8%; Score 1256; DB 4; Length 228;
 Best Local Similarity 100.0%; Pred. No. 1.8e-107;
 Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

16 SPPLDNKTPSLCPGHPGLPGTGHGSGCLPGRDRDGPAGPGEKGGGRLPG 75
 |||||||
 1 SPPLDNKTPSLCPGHPGLPGTGHGSGCLPGRDRDGPAGPGEKGGGRLPG 60
 76 PRGDDPGRGEGAPAGTGTGAGGECVPPRASFSAKRSRVPSPDAPLPFDRVLVNEQGH 135
 |||||||
 61 PRGDDPGRGEGAPAGTGTGAGGECVPPRASFSAKRSRVPSPDAPLPFDRVLVNEQGH 120
 136 YDAVTKGFTCOVPGVYFAVHATVYRASLQFDLVKNGESIASFFQFGGMPKPAISLSCGA 195
 |||||||
 121 YDAVTKGFTCOVPGVYFAVHATVYRASLQFDLVKNGESIASFFQFGGMPKPAISLSCGA 180
 196 MVRLEPEDQVWVOVGVDYIGIYASIKTSTSGFLVYSDMHSSPVFA 243
 |||||||
 181 MVRLEPEDQVWVOVGVDYIGIYASIKTSTSGFLVYSDMHSSPVFA 228

RESULT 6

US-09-336-536-11
 : Sequence 11, Application US/09336536
 : Patent No. 6406884
 : GENERAL INFORMATION:
 : APPLICANT: Leiby, K.
 : APPLICANT: McKay, C.
 : APPLICANT: Bossone, S.
 : TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
 : FILE REFERENCE: 7853-144
 : CURRENT APPLICATION NUMBER: US/09/336, 536
 : CURRENT FILING DATE: 1999-06-18
 : NUMBER OF SEQ ID NOS: 75
 : SOFTWARE: Patentln Ver. 2.0
 : SEQ ID NO 11
 : LENGTH: 228
 : TYPE: PRT
 : ORGANISM: Mus musculus
 US-09-336-536-11

Query Match 90.6%; Score 1200; DB 4; Length 228;
 Best Local Similarity 94.7%; Pred. No. 2.4e-102;
 Matches 216; Conservative 5; Mismatches 7; Indels 0; Gaps 0;

16 SPPLDNKTPSLCPGHPGLPGTGHGSGCLPGRDRDGPAGPGEKGGGRLPG 75
 |||||||
 1 SPPLDNKTPSLCPGHPGLPGTGHGSGCLPGRDRDGPAGPGEKGGGRLPG 60
 76 PRGDDPGRGEGAPAGTGTGAGGECVPPRASFSAKRSRVPSPDAPLPFDRVLVNEQGH 135
 |||||||
 61 PRGDDPGRGEGAPAGTGTGAGGECVPPRASFSAKRSRVPSPDAPLPFDRVLVNEQGH 120
 136 YDAVTKGFTCOVPGVYFAVHATVYRASLQFDLVKNGESIASFFQFGGMPKPAISLSCGA 195
 |||||||
 121 YDAVTKGFTCOVPGVYFAVHATVYRASLQFDLVKNGESIASFFQFGGMPKPAISLSCGA 180
 196 MVRLEPEDQVWVOVGVDYIGIYASIKTSTSGFLVYSDMHSSPVFA 243
 |||||||
 181 MVRLEPEDQVWVOVGVDYIGIYASIKTSTSGFLVYSDMHSSPVFA 228

RESULT 7

US-09-336-536-7
 : Sequence 7, Application US/09336536

Patent No. 6406884
 : GENERAL INFORMATION:
 : APPLICANT: Leiby, K.
 : APPLICANT: McKay, C.
 : APPLICANT: Bossone, S.
 : TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
 : FILE REFERENCE: 7853-144
 : CURRENT APPLICATION NUMBER: US/09/336, 536
 : CURRENT FILING DATE: 1999-06-18
 : NUMBER OF SEQ ID NOS: 75
 : SOFTWARE: Patentln Ver. 2.0
 : SEQ ID NO 7
 : LENGTH: 128
 : TYPE: PRT
 : ORGANISM: Homo sapiens
 US-09-336-536-7

Query Match 50.9%; Score 675; DB 4; Length 128;
 Best Local Similarity 100.0%; Pred. No. 1.2e-54;
 Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

105 AFSARSRSESRVPPSDAPLPFDRVLVNEQGHYDAVTKGFTCOVPGVYFAVHATVYRASL 164
 |||||||
 1 AFSARSRSESRVPPSDAPLPFDRVLVNEQGHYDAVTKGFTCOVPGVYFAVHATVYRASL 60
 165 QFDLVKNGESIASFFQFGGMPKPAISLSCGAMVRLEPEDQVWVOVGVDYIGIYASIKTD 224
 |||||||
 61 QFDLVKNGESIASFFQFGGMPKPAISLSCGAMVRLEPEDQVWVOVGVDYIGIYASIKTD 120
 225 STFSGFLV 232
 |||||||
 121 STFSGFLV 128

RESULT 8

US-09-336-536-14
 : Sequence 14, Application US/09336536
 : Patent No. 6406884
 : GENERAL INFORMATION:
 : APPLICANT: Leiby, K.
 : APPLICANT: McKay, C.
 : APPLICANT: Bossone, S.
 : TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
 : FILE REFERENCE: 7853-144
 : CURRENT APPLICATION NUMBER: US/09/336, 536
 : CURRENT FILING DATE: 1999-06-18
 : NUMBER OF SEQ ID NOS: 75
 : SOFTWARE: Patentln Ver. 2.0
 : SEQ ID NO 14
 : LENGTH: 128
 : TYPE: PRT
 : ORGANISM: Mus musculus
 US-09-336-536-14

Query Match 49.1%; Score 650; DB 4; Length 128;
 Best Local Similarity 94.5%; Pred. No. 2.3e-52;
 Matches 121; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

105 AFSARSRSESRVPPSDAPLPFDRVLVNEQGHYDAVTKGFTCOVPGVYFAVHATVYRASL 164
 |||||||
 1 AFSARSRSESRVPPSDAPLPFDRVLVNEQGHYDAVTKGFTCOVPGVYFAVHATVYRASL 60
 165 QFDLVKNGESIASFFQFGGMPKPAISLSCGAMVRLEPEDQVWVOVGVDYIGIYASIKTD 224
 |||||||
 61 QFDLVKNGESIASFFQFGGMPKPAISLSCGAMVRLEPEDQVWVOVGVDYIGIYASIKTD 120
 225 STFSGFLV 232
 |||||||
 121 STFSGFLV 128

RESULT 9

US-08-463-911-2

Sequence 2, Application US/08463911
Patent No. 5869330
GENERAL INFORMATION:
APPLICANT: Scherer, Philipp E.
APPLICANT: Lodish, Harvey F.
TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millita Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION NUMBER: US/08/463,911
APPLICATION NUMBER: 530
CLASSIFICATION: 530
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH195-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 247 amino acids
TYPE: amino acid
TOPOLOGY: Linear
MOLECULE TYPE: protein
US-08-463-911-2

Query Match 32.7%; Score 433.5; DB 2; Length 247;
Best Local Similarity 40.6%; Pred. No. 3,4e-32;
Matches 102; Conservative 32; Mismatches 92; Indels 25; Gaps 8;

QY 1 MRPLVLLLL-GLAAGSPPLDNNKIPSLCPGHPG-----LPGTGHHGSGGLPGRDGRD 54
DB 4 LQALLFLILPSHAEDDVTTEELAPALVPPPKCTCAGMMAGIGPHGHNCTPGRDGRD- 62
QY 55 RDGAPGAPGEGEGRGRLPGPRGDP-----PGPRGEGAPGPTGPRGECSEVPPRSA 105
DB 63 -----GTPGEGEGGAGLGLGPKGTGDTGVTGAGPRGPTGPRGEGCEAAVMYRSA 117
QY 106 FSAKRSRVPPEPSDAPLPEDRVLVNEQGHYDAVTGKFTCOVPGVYFAVAHATYRASLQ 165
DB 118 FSV-GLETRVTV-PNVIRFTKIFYNOONHYDSTGKFCYNIPGLYFSYHITYYMDVK 175
QY 166 FDLVKNESIASFQFGGKMPKPSLSGAMVRLEPEDQVWVY-GVGDYIGIYASIKTD 224
DB 176 VSLPKKRAVLEFYDYQOE-KNVDAQSGSVLHLEVDQWLVLYGQGDHGLYADVND 234
QY 225 STEGFLVYSD 235
DB 235 STEGFLVYHD 245

RESULT 10
US-09-140-804-8
Sequence 8, Application US/09140804
Patent No. 6197930
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Humes, Jacqueline M.
TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
FILE REFERENCE: 97-49

CURRENT APPLICATION NUMBER: US/09/140,804
CURRENT FILING DATE: 1998-08-26
EARLIER APPLICATION NUMBER: 60/056,983
EARLIER FILING DATE: 1997-08-26
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 8
LENGTH: 247
TYPE: PRT
ORGANISM: Homo sapiens
US-09-140-804-8

Query Match 32.3%; Score 428.5; DB 4; Length 247;
Best Local Similarity 40.2%; Pred. No. 9,7e-32;
Matches 101; Conservative 32; Mismatches 93; Indels 25; Gaps 8;

QY 1 MRPLVLLLL-GLAAGSPPLDNNKIPSLCPGHPG-----LPGTGHHGSGGLPGRDGRD 54
DB 4 LQALLFLILPSHAEDDVTTEELAPALVPPPKCTCAGMMAGIGPHGHNCTPGRDGRD- 62
QY 55 RDGAPGAPGEGEGRGRLPGPRGDP-----PGPRGEGAPGPTGPRGECSEVPPRSA 105
DB 63 -----GTPGEGEGGAGLGLGPKGTGDTGVTGAGPRGPTGPRGEGCEAAVMYRSA 117
QY 106 FSAKRSRVPPEPSDAPLPEDRVLVNEQGHYDAVTGKFTCOVPGVYFAVAHATYRASLQ 165
DB 118 FSV-GLETRVTV-PNVIRFTKIFYNOONHYDSTGKFCYNIPGLYFSYHITYYMDVK 175
QY 166 FDLVKNESIASFQFGGKMPKPSLSGAMVRLEPEDQVWVY-GVGDYIGIYASIKTD 224
DB 176 VSLPKKRAVLEFYDYQOE-KNVDAQSGSVLHLEVDQWLVLYGQGDHGLYADVND 234
QY 225 STEGFLVYSD 235
DB 235 STEGFLVYHD 245

RESULT 11
US-09-118-408-3
Sequence 3, Application US/09118408A
Patent No. 6265544
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
FILE REFERENCE: 97-30
CURRENT APPLICATION NUMBER: US/09/118,408A
CURRENT FILING DATE: 1998-07-17
EARLIER APPLICATION NUMBER: 60/053,154
EARLIER FILING DATE: 1997-07-18
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 247
TYPE: PRT
ORGANISM: Homo sapiens
US-09-118-408-3

Query Match 32.3%; Score 428.5; DB 4; Length 247;
Best Local Similarity 40.2%; Pred. No. 9,7e-32;
Matches 101; Conservative 32; Mismatches 93; Indels 25; Gaps 8;

QY 1 MRPLVLLLL-GLAAGSPPLDNNKIPSLCPGHPG-----LPGTGHHGSGGLPGRDGRD 54
DB 4 LQALLFLILPSHAEDDVTTEELAPALVPPPKCTCAGMMAGIGPHGHNCTPGRDGRD- 62
QY 55 RDGAPGAPGEGEGRGRLPGPRGDP-----PGPRGEGAPGPTGPRGECSEVPPRSA 105
DB 63 -----GTPGEGEGGAGLGLGPKGTGDTGVTGAGPRGPTGPRGEGCEAAVMYRSA 117
QY 106 FSAKRSRVPPEPSDAPLPEDRVLVNEQGHYDAVTGKFTCOVPGVYFAVAHATYRASLQ 165
DB 118 FSV-GLETRVTV-PNVIRFTKIFYNOONHYDSTGKFCYNIPGLYFSYHITYYMDVK 175

QY 166 FDLVKNGESIASFEFGGMPKRPASLSGAMVRLEREDQVWVYV-GVGDYIGIYASIKTD 224
Db 176 VSLFKKDKAVLFTYDQYOE-KNVYDASGSVLLHLLEVGDQVWLQVYGGDHNGILYADNVND 234
QY 225 STFSGFLVYSD 235
Db 235 STFTGFLLYHD 245

RESULT 12

US-09-506-855-3
Sequence 3, Application US/09506855
Patent No. 6448221
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Lasser, Gerald W.
APPLICANT: Bishop, Paul D.
TITLE OF INVENTION: INHIBITORS FOR USE IN HEMOSTASIS AND
FILE REFERENCE: 99-12
CURRENT APPLICATION NUMBER: US/09/506,855
CURRENT FILING DATE: 2000-02-17
NUMBER OF SEQ ID NOS: 50
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 247
TYPE: PRP
ORGANISM: Homo sapiens
US-09-506-855-3

Query Match 32.3%; Score 428.5; DB 4; Length 247;
Best Local Similarity 40.2%; Pred. No. 9, 7e-32;
Matches 101; Conservative 32; Mismatches 93; Indels 25; Gaps 8;

QY 1 MRPLVLLLL-GLAASPLDNNKIRSLCPGHPG-----LPGRGHHGSGCLGRDGRDG 54
Db 4 LQALLLELLPSHAEDQVTTTELALVLPKGTGAGMAGIPGHPGHTGTRGDRGD- 62
QY 55 RDGAPGAPGEGKGRGRLPGPRGDPGPGGAGPAPG-----TGPAGGCSVPRPSA 105
Db 63 -----CTPEKKGKAGLGRGEGTGDVGMGAEPRGPGTGRKGRGEGEAYWRSA 117
QY 106 FSAKRSSESVPPSDAPLPEDRVLVNEQGHYDAVTGKTCQVPGVYVFAVHATVYRASIQ 165
Db 118 FSV-GLETRYVAP-NVPIRFTKIFYNQNHNDOSTGKFCNIEGLYFVSHITVYMKDVK 175
QY 166 FDLVKNGESIASFEFGGMPKRPASLSGAMVRLEREDQVWVYV-GVGDYIGIYASIKTD 224
Db 176 VSLFKKDKAVLFTYDQYOE-KNVYDASGSVLLHLLEVGDQVWLQVYGGDHNGILYADNVND 234
QY 225 STFSGFLVYSD 235
Db 235 STFTGFLLYHD 245

RESULT 13

US-08-463-911-7
Sequence 7, Application US/08463911
Patent No. 5869330
GENERAL INFORMATION:
APPLICANT: Scherer, Philipp E.
APPLICANT: Lodish, Harvey F.
TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463,911
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WHI95-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 244 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-463-911-7

Query Match 32.0%; Score 424; DB 2; Length 244;
Best Local Similarity 40.5%; Pred. No. 2, 5e-31;
Matches 106; Conservative 26; Mismatches 72; Indels 58; Gaps 10;

QY 6 VLLLLGLAASPLDNNKIRSLCPGHPG-----PG-----LPGRGHHGSG 44
Db 7 VLLLLGLAASPLDNNKIRSLCPGHPG-----PGHQDETITGPGVLLRLPKGACTGMAIGIPGHPG 50
QY 45 GLPGRGGRDGRDGPAPGPEKGGRRPLGPRGCD-----PPRPGAGPAGTGRA 95
Db 51 GAGRGDRD-----CTPEKKGKAGLGRGEGTGDVGMGAEPRGPGTGRKGRG 104
QY 96 GECSVPRSAFSAKRSSESVPPSDAPLPEDRVLVNEQGHYDAVTGKTCQVPGVYVFAV 155
Db 105 GEGAVYYSASFV-GLETYYTIP-NMPIRFTKIFYNQNHNDOSTGKFCNIEGLYFV 162
QY 156 HATVYRASLQFDLVKNGES-IASFFOFGGMPKRPASLSGAMVRLEREDQVWVYV-GVGD 213
Db 163 HTVYMKDVKVSLFKKDKAMLFTYDQYQENNDQA--SGSVLLHLLEVGDQVWLQVYGGEGE 220
QY 214 YIGIYASIKTDSTFSGFLVYSD 235
Db 221 RNLGYADNNDSTFTGFLLYHD 242

RESULT 14

US-09-140-804-3
Sequence 3, Application US/09140804
Patent No. 6197930
GENERAL INFORMATION:
APPLICANT: Sheppard, Paul O.
APPLICANT: Humes, Jacqueline M.
TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
FILE REFERENCE: 97-49
CURRENT APPLICATION NUMBER: US/09/140,804
CURRENT FILING DATE: 1998-08-26
EARLIER APPLICATION NUMBER: 60/056,983
EARLIER FILING DATE: 1997-08-26
NUMBER OF SEQ ID NOS: 47
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 3
LENGTH: 244
TYPE: PRP
ORGANISM: Homo sapiens
US-09-140-804-3

Query Match 32.0%; Score 424; DB 4; Length 244;
Best Local Similarity 40.5%; Pred. No. 2, 5e-31;
Matches 106; Conservative 26; Mismatches 72; Indels 58; Gaps 10;

```
OY 6 VLLLLGLAAGSPPLDNDKIPSLCPGH-----PG-----LPGTPGHHGSQ 44
    |||||
    7 VLLLLAL-----PGHDOETTTGPGVLLPLPKGACTGMMAGIPGHPGN 50
OY 45 GLPGRGRDGRDGAAPGAPGEGKPGGLPGPRGD-----PGPRGEGAPGAPTPGPA 95
    |||||
    51 GARGRGRD-----GTPGEGKGGDPGLGPKGDIGETGVPAGEGPRGPGIOGRKGP 104
OY 96 GECSVPPRASFSAKRSRPPPSDAPLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAY 155
    |||||
    105 GEGAVYVRSASFV-GLETTYTIP-NMPIRTKIFYNQNNHYDSTGKFNHCPGLYFAY 162
OY 156 HATVYRASLQDFLVKNGES-IASFQPFGGMPKPAISGAMVRLPEPDQVWVQV-GVGD 213
    |||||
    163 HITVYMKDVAVSLFKKDKAMLFYYDQYQENNVDA--SGSVLLHLEVGDQVWLQVYGEGE 220
OY 214 YIGIVASIKTDSFSGFLVYSD 235
    |||||
    221 RKGIVADNDNDSTFTGFLYHD 242
Db
```

RESULT 15

```
US-09-336-536-20
; Sequence 20, Application US/09336536
; Patent No. 640884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 20
; LENGTH: 244
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-536-20
```

Query Match 32.0%; Score 424; DB 4; Length 244;

Best Local Similarity 40.5%; Pred. No. 2.5e-31; Matches 106; Conservative 26; Mismatches 72; Indels 58; Gaps 10;

```
OY 6 VLLLLGLAAGSPPLDNDKIPSLCPGH-----PG-----LPGTPGHHGSQ 44
    |||||
    7 VLLLLAL-----PGHDOETTTGPGVLLPLPKGACTGMMAGIPGHPGN 50
OY 45 GLPGRGRDGRDGAAPGAPGEGKPGGLPGPRGD-----PGPRGEGAPGAPTPGPA 95
    |||||
    51 GARGRGRD-----GTPGEGKGGDPGLGPKGDIGETGVPAGEGPRGPGIOGRKGP 104
OY 96 GECSVPPRASFSAKRSRPPPSDAPLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAY 155
    |||||
    105 GEGAVYVRSASFV-GLETTYTIP-NMPIRTKIFYNQNNHYDSTGKFNHCPGLYFAY 162
OY 156 HATVYRASLQDFLVKNGES-IASFQPFGGMPKPAISGAMVRLPEPDQVWVQV-GVGD 213
    |||||
    163 HITVYMKDVAVSLFKKDKAMLFYYDQYQENNVDA--SGSVLLHLEVGDQVWLQVYGEGE 220
OY 214 YIGIVASIKTDSFSGFLVYSD 235
    |||||
    221 RKGIVADNDNDSTFTGFLYHD 242
Db
```

Search completed: June 18, 2003, 14:54:23
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OW protein - protein search, using sw model

Run on: June 18, 2003, 14:54:28 ; Search time 23 Seconds

(without alignments)
438.206 Million cell updates/sec

Title: US-09-943-851A-42

Perfect score: 243
Sequence: 1 MRPLVLVLGLAGSPPLD.....DSTFSGELVYSDMHSSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 112892 seqs, 41476328 residues

Word size : 0

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : SwissProt_40.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	100.0	243	1 CQT5_HUMAN	Q9BXJ0 homo sapien
2	11	4.5	289	1 CQT7_HUMAN	Q9BXJ2 homo sapien
3	11	4.5	369	1 PSPD_BOVIN	P35246 bos taurus
4	11	4.5	374	1 PSPD_MOUSE	P50404 mus musculu
5	11	4.5	374	1 PSPD_RAT	P35248 rattus norv
6	11	4.5	375	1 PSPD_HUMAN	P35247 homo sapien
7	11	4.5	1355	1 CA21_RANCA	Q42350 rana catesb
8	11	4.5	1670	1 CA34_HUMAN	Q01955 homo sapien
9	11	4.5	2944	1 CA17_HUMAN	Q02388 homo sapien
10	10	4.1	210	1 CAC2_HAECO	P16252 haemochus
11	10	4.1	247	1 PSPA_CAVPO	P50403 cavia porce
12	10	4.1	247	1 PSPA_RABIT	P12842 oryctolagus
13	10	4.1	248	1 PSPA_CANFA	P07714 homo sapien
14	10	4.1	248	1 PSPA_HUMAN	P06908 canis fami
15	10	4.1	248	1 PSPA_MOUSE	P35242 mus musculu
16	10	4.1	248	1 PSPA_RAT	P08427 rattus norv
17	10	4.1	249	1 PSPA_PIG	P49874 sus scrofa
18	10	4.1	526	1 CA21_RABIT	Q28668 oryctolagus
19	10	4.1	547	1 CARL_EPHMU	P18856 ephedalia m
20	10	4.1	1049	1 CA13_BOVIN	P04258 bos taurus
21	10	4.1	1464	1 CA13_MOUSE	P08121 mus musculu
22	10	4.1	1466	1 CA13_HUMAN	P02461 homo sapien
23	9	3.7	265	1 CTRC_NEIMA	P57012 neisseria m
24	9	3.7	336	1 CA13_RAT	P13941 rattus norv
25	9	3.7	695	1 APP2_MOUSE	P06335 mus musculu
26	9	3.7	765	1 APP2_RAT	P15943 rattus norv
27	9	3.7	1364	1 CA21_BOVIN	P02465 bos taurus
28	9	3.7	1366	1 CA21_CANFA	Q46392 canis fami
29	9	3.7	1402	1 IF4G_RABIT	P41110 oryctolagus
30	9	3.7	1669	1 CA14_MOUSE	P02463 mus musculu
31	8	3.3	28	1 C10C_RAT	P31722 rattus norv
32	8	3.3	170	1 FA39_HUMAN	P49913 homo sapien
33	8	3.3	193	1 CERR_HUMAN	P23435 homo sapien

34	8	3.3	193	1 CERR_MOUSE	Q91711 mus musculu
35	8	3.3	201	1 CERR_HUMAN	Q9NCU7 homo sapien
36	8	3.3	224	1 CERR_RAT	P98087 rattus norv
37	8	3.3	244	1 APM1_HUMAN	Q15848 homo sapien
38	8	3.3	247	1 APM1_MOUSE	P06394 mus musculu
39	8	3.3	301	1 CC02_MOUSE	P17656 caenorhabdl
40	8	3.3	306	1 CC40_CAEEL	P34804 caenorhabdl
41	8	3.3	316	1 CC12_CAEEL	P20630 caenorhabdl
42	8	3.3	316	1 CC13_CAEEL	P20631 caenorhabdl
43	8	3.3	329	1 CQT4_HUMAN	Q9BXJ3 homo sapien
44	8	3.3	356	1 VP36_CANFA	P49256 canis fami
45	8	3.3	358	1 VP36_MOUSE	Q9DBH5 mus musculu

ALIGNMENTS

RESULT 1
ID CQT5_HUMAN STANDARD: PRT; 243 AA.
AC Q9BXJ0; Q9BFX4; 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Complement-clq tumor necrosis factor-related protein 5 precursor.
GN C1QTNF5 OR CTRP5.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Shepard P.O., Humes J.M.;
RT "Homo sapiens complement-clq tumor necrosis factor-related protein.";
RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 25-243 FROM N.A.
RC TISSUE=uterus;
RA Ottenwaelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF329841; AAK17965.1; -
CC EMBL; AL110261; CAB53702.1; -
CC Genew; HGNC:14344; C1QTNF5.
CC Interpro: IPR001073; C1q.
CC Interpro: IPR000087; Collagen.
CC Pfam; PF00386; C1q; 1.
CC Pfam; PF01391; Collagen; 1.
CC PRINTS; PR00007; COMPLEMENTC1Q.
CC SMART; SM00110; C1Q; 1. FALSE_NEG.
DR PROSITE; PS01113; C1Q; FALSE_NEG.
KW Collagen; Signal.
FT CHAIN 1 15 POTENTIAL.
FT 16 243 COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT 243 RELATED PROTEIN 5.
FT DOMAIN 30 95 COLLAGEN-LIKE.
FT 243 C1Q.
SQ SEQUENCE 243 AA; 25298 MW; 7CCDA65CDA7EB784 CRC64;
Query Match 100.0%; Score 243; DB 1; Length 243;
Best Local Similarity 100.0%; Pred. No. 1,9e+193;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

[illegible]

RESULT 2	COT7_HUMAN		STANDARD:	PRT: 289 AA.
ID	COT7_HUMAN	STANDARD:	PRT: 289 AA.	
AC	Q9BXJ2;			
DT	15-JUN-2002 (Rel. 41, Created)			
DT	15-JUN-2002 (Rel. 41, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Complement-c1q tumor necrosis factor-related protein 7 precursor.			
GN	CIQTNF7 OR CTRP7.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.			
OX	NCBI_Taxid=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Piddington C.S., Sheppard P.O., Bishop P., Lasser G.W.;			
RT	"Homo sapiens complement-c1q tumor necrosis factor-related protein.";			
RL	Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE-Testis;			
RA	Straussberg R.;			
RL	Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.			
CC	-1- SIMILARITY: CONTAINS 1 COLLAGENOUS DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 1 C1Q DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation			
CC	at the European Bioinformatics Institute. There are no restrictions on			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see http://www.isb-slb.ch/announce/			
CC	or send an email to license@isb-slb.ch).			
CC	-----			
DR	EMBL; AF329839; AAK17963.1; -			
DR	EMBL; BC022187; AAK22187.1; -			
DR	Genew; HGNC:14342; CIQTNF7.			
DR	InterPro; IPR001073; C1q.			
DR	InterPro; IPR000087; Collagen.			
DR	Pfam; PF00386; C1q; 1.			
DR	Pfam; PF01391; Collagen; 2.			
DR	PRINTS; PR00007; COMPLEMENTC1Q.			
DR	SMART; SM00110; C1Q; 1.			
DR	PROSITE; PS0113; C1Q; 1.			
KM	Collagen; Signal.			
FT	SIGNAL	1	16	POTENTIAL.
FT	CHAIN	17	289	COMPLEMENT-C1Q TUMOR NECROSIS FACTOR-
FT	DOMAIN	38	139	RELATED PROTEIN 7.
FT	DOMAIN	141	276	COLLAGEN-LIKE.
FT	SEQUENCE	289 AA;	30683 MW;	A61609FE6D26946 CRC64;

Query Match	4.58;	Score 11;	DB 1;	Length 289;
Best Local Similarity	100.08;	Pred. No. 0.06;		
Matches 11; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

```
Qy      45 GLPRDGRDGR 55
         |||||
Db      59 GLPRDGRDGR 69
```

RESULT 3

ID	PSPD_BOVIN	STANDARD:	PRT:	369 AA.
AC	p352d6:			
DT	01-FEB-1994 (Rel. 28, Created)			
DT	01-FEB-1994 (Rel. 28, Last sequence update)			
DT	15-JUL-1999 (Rel. 38, Last annotation update)			
DE	Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D)			
GN	SEPPD OR SFPD4			
OS	Bos taurus (Bovine).			
OC	Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;			
OC	Bovidae; Bovinae; Bos.			
OX	NCBI_TaxID=9913;			

RP SEQUENCE FROM N.A., AND SEQUENCE OF 208-247.

RX MEDLINE=93170856; PubMed=8436402;

RT "Structural similarity between bovine conglutinin and bovine lung

synthesis of conglutinin.";

CC -1- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED

EXTENT OTHER ALPHA-GLUCOSYL MOIETIES. IT COULD PARTICIPATE IN THE

-1- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.
CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%

CC CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
CC PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGEN

CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC HYDROPHOBIC PROTEINS (SP-B AND SP-C).

CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN

CC
CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation in
CC Hinxton, UK.

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial purposes is prohibited. For more information, see <https://creativecommons.org/licenses/by/4.0/>.

CC or send an email to license@isb-sib.ch.

DR EMBL; X75911; CAA53510.1; -.

DR HSSP; P35247; 1B08.

DR InterPro; IPR001304; Lectin_C.

DR Pfam; PF01391; Collagen; 2.

```
DR      SMN1; CTYPE_LECTIN_1; 1.
DR      PROSITE; PS00615; C_TYPE_LECTIN_1; 1.
```

glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation; KW

KW	SIGNAL; Declin; collagen; repeat; coll. coll.
FT	SIGNAL 1 20 BY SIMILARITY.

FT			D.
ET	CHAIN	21	509
			POLYMONARI SUBFACIANT ASSOCIATED FOLIAN

FT	217	248	COILED COIL (POTENTIAL).
FT	46	216	COLLAGEN-LIKE.
FT	217	248	COILED COIL (POTENTIAL).
FT	46	216	COLLAGEN-LIKE.

FT	DOMAIN	273	369	C-TYPE LECTIN (SHORT FORM).
FT	DISULFID	275	367	BY SIMILARITY.

FT	DISULFID	345	359	BY SIMILARITY.
FT	CARBOHYD	90	90	N-LINKED (GLCNAC.) (POTENTIAL).
FT	MOD_RES	78	78	HYDROXYLATION (BY SIMILARITY).
FT	MOD_RES	87	87	HYDROXYLATION (BY SIMILARITY).
FT	MOD_RES	96	96	HYDROXYLATION (BY SIMILARITY).
FT	MOD_RES	99	99	HYDROXYLATION (BY SIMILARITY).
FT	MOD_RES	165	165	HYDROXYLATION (BY SIMILARITY).
FT	MOD_RES	171	171	HYDROXYLATION (BY SIMILARITY).
SO	SEQUENCE	369 AA:	37361 MW; 07D8B24E0AB2E3 CRC64;	
Query Match				
Best Local Similarity		100.0%;	Score 11; DB 1; Length 369;	
Matches 11; Conservative		0;	Pred. No. 0.073; Mismatches 0; Indels 0; Gaps 0;	
QY	45 GLPGRDGRGR 55			
Db	46 GLPGRDGRGR 56			
RESULT 4				
PSPD_MOUSE	STANDARD;	PRT;	374 AA.	
AC	PSPD_MOUSE			
AC	P50404;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-JUN-2002 (Rel. 41, Last annotation update)			
DE	Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).			
GN	SEPPD OR SFTPD.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6 X CBA; TISSUE=Lung;			
RX	MEDLINE=96094460; PubMed=7499852;			
RA	Moswani M., White R.A., Guo N., Dowler L.L., Tauber A.I., Sastry K.N.;			
RT	"Mouse surfactant protein-D. cDNA cloning, characterization, and gene			
RL	localization to chromosome 14.";			
RL	J. Immunol. 155:5671-5677(1995).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=129/Sv;			
RX	MEDLINE=99244602; PubMed=10226052;			
RA	Lawson P.R., Perkins V.C., Holmskov U., Reid K.B.;			
RT	"Genomic organization of the mouse gene for lung surfactant protein			
RT	D.";			
RL	Am. J. Respir. Cell Mol. Biol. 20:953-963(1999).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	Fisher J.H., Shettelyevich V.V.;			
RT	"Surfactant protein-D regulates surfactant phospholipid homeostasis in			
RT	vivo.";			
RL	Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RA	Strausberg R.;			
RL	Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.			
CC	-1- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED			
CC	MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER			
CC	EXTENT OTHER ALPHA-GLUCOSYL MOIETIES. IT COULD PARTICIPATE IN THE			
CC	EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.			
CC	-1- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.			
CC	-1- SUBCELLULAR LOCATION: Extracellular.			
CC	-1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%			
CC	PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,			
CC	CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-B AND SP-C).			
CC	HYDROPHOBIC PROTEINS (SP-A AND SP-D) AND 2 SMALL			
CC	-1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.			
CC	-1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			

```

CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@lsb-sib.ch).
```

```

CC EMBL: LA0156; AAA92021.1;
CC
CC EMBL: AF047742; AAD31380.1;
CC DR EMBL: AF047741; AAD31380.1; JOINED.
CC DR EMBL: AF192134; AAF15277.1;
CC DR EMBL: BC003705; AAH03705.1;
CC DR HSSP: P35247; 1808.
CC MGD: MGI:109515; Sfcpd.
CC InterPro: IPR000087; Collagen.
CC InterPro: IPR001304; Lectin_C.
CC Pfam: PF00059; Lectin_C; 1.
CC DR Pfam: PF01391; Collagen; 3.
CC DR SMART: SMO0034; CLECT; 1.
CC DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
CC DR PROSITE: PS50041; C-TYPE_LLECTIN_2; 1.
CC KW Glycoprotein; Calcium; Surface Film; Gaseous exchange; Hydroxylation;
CC Signal; Lectin; Collagen; Repeat; Coiled coil.
CC FT SIGNAL 1 19
CC FT CHAIN 20 374
CC FT D.
CC FT DOMAIN 45 221
CC FT COLLAGEN-LIKE.
CC FT DOMAIN 222 253
CC FT COILED COIL (POTENTIAL).
CC FT DOMAIN 278 374
CC FT C-TYPE LECTIN (SHORT FORM).
CC FT DISULFID 280 372
CC FT BY SIMILARITY.
CC FT DISULFID 350 364
CC FT BY SIMILARITY.
CC FT CARBOHYD 89
CC FT N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SO SEQUENCE 374 AA; 37688 MW; FE034261263f43fe4 CMC64;

Query Match 4.5%; Score 11; DB 1; Length 374;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRGR 55
DB 45 GLPGRDGRGR 55
|||||
|

RESULT 5
PSPD_RAT STANDARD: PRT: 374 AA.
AC P35248;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D)
DE (CP4).
GN SFMPD OR SFPP4.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OC NCBI TaxID=10116;
RN [1]
RN SEQUENCE FROM N.A., AND SEQUENCE OF 20-33.
RP TISSUE=Lung;
RC MEDLINE=92112913; PubMed=1370483;
RA Shimizu H., Fisher J.H., Papst P., Benson B., Lau K., Mason R.J.,
RA Voelker D.R.;
RA "Primary structure of rat pulmonary surfactant protein D. cDNA and
RA deduced amino acid sequence.";
RL J. Biol. Chem. 267:1853-1857(1992).
RP [2]
RP SEQUENCE OF 73-95 AND 153-180.
RC TISSUE=Lung;
RA MEDLINE=90001186; PubMed=2675969;
RA Person A., Chang D., Rust K., Moxley M., Longmore W., Crouch E.;
RA "Purification and biochemical characterization of CP4 (SP-D), a
RA collagenous surfactant-associated protein.";
RL Biochemistry 28:6361-6367(1989).
```

CC -1- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED
 CC MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER
 CC EXTENT OTHER ALPHA-GLUCOSYL MOETIES. IT COULD PARTICIPATE IN THE
 CC EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.
 CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
 CC PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
 CC CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
 CC HYDROPHOBIC PROTEINS (SP-B AND SP-C).
 CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M81231; AAA42170.1; -
 CC PIR: A42046; A42046.
 CC HSSP: P35247; 1B08.
 CC InterPro: IPR000087; Collagen.
 CC InterPro: IPR001304; Lectin_C.
 CC Pfam: PF00059; Lectin_C; 1.
 CC Pfam: PF01391; Collagen; 3.
 CC SMART: SM00034; CLECT; 1.
 CC PROSITE: PS00615; C-TYPE LECTIN_1; 1.
 CC PROSITE: PS50041; C-TYPE LECTIN_2; 1.
 CC Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
 KW Signal; Lectin; Collagen; Repeat; Collid coll.
 FT CHAIN 1 19
 FT 20 374 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
 FT D.
 FT DOMAIN 45 221 COLLAGEN-LIKE.
 FT DOMAIN 222 253 COILED COIL (POTENTIAL).
 FT 278 374 C-TYPE LECTIN (SHORT FORM).
 FT DISULFID 280 372 BY SIMILARITY.
 FT DISULFID 350 364 BY SIMILARITY.
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .).
 FT MOD_RES 77 77 HYDROXYLATION.
 FT MOD_RES 86 86 HYDROXYLATION.
 FT MOD_RES 95 95 HYDROXYLATION.
 FT MOD_RES 98 98 HYDROXYLATION.
 FT MOD_RES 170 170 HYDROXYLATION.
 FT MOD_RES 176 176 HYDROXYLATION.
 FT CONFLICT 89 89 N -> E (IN REF. 2).
 FT CONFLICT 164 164 K -> C (IN REF. 2).
 SO SEQUENCE 374 AA; 37561 MW; D82B85E39D4A3C CRC64;
 Query Match 4.5%; Score 11; DB 1; Length 374;
 Best Local Similarity 100.0%; Pred. No. 0.074;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 45 GLPGRDGRGR 55
 Db 45 GLPGRDGRGR 55
 RESULT 6
 PSPD_HUMAN STANDARD: PRT: 375 AA.
 ID PSPD_HUMAN P35247;
 AC P35247;
 DT 01-FEB-1994 (Rel. 28, Created)
 DT 01-FEB-1994 (Rel. 28, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pulmonary surfactant-associated protein D precursor (SP-D) (PSP-D).
 GN SFTPD OR SFTPD OR PSPD.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=9155122; PubMed=8428971;
 RA Crouch E., Rust K., Velle R., Donis-Keller H., Grosso L.;
 RT "Genomic organization of human surfactant protein D (SP-D). SP-D is
 RT encoded on chromosome 10q22.2-23.1.";
 RL J. Biol. Chem. 268:2976-2983(1993).
 RN [2]
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 214-243.
 RC TISSUE=Lung, and Amniotic fluid;
 RX MEDLINE=92322003; PubMed=1339284;
 RA Lu J., Willis A.C., Reid K.B.M.;
 RT "Purification, characterization and cDNA cloning of human lung
 RT surfactant protein D.";
 RL Biochem. J. 284:795-802(1992).
 RN [3]
 RP SEQUENCE OF 60-375 FROM N.A., AND PARTIAL SEQUENCE.
 RC TISSUE=Lung;
 RX MEDLINE=91378578; PubMed=1898081;
 RA Rust K., Grosso L., Zhang V., Chang D., Petsson A., Longmore W.,
 RA Cai G.-Z., Crouch E.;
 RT "Human surfactant protein D: SP-D contains a C-type lectin
 RT carbohydrate recognition domain.";
 RL Arch. Biochem. Biophys. 290:116-126(1991).
 RN [4]
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS).
 RX MEDLINE=99197291; PubMed=10368295;
 RA Hakansson K., Lim N.K., Hoppe H.-J., Reid K.B.M.;
 RT "Crystal structure of the trimeric alpha-helical coiled-coil and the
 RT three lectin domains of human lung surfactant protein D.";
 RL Structure 7:255-264(1999).
 CC -1- FUNCTION: CONTRIBUTES TO THE LUNG'S DEFENSE AGAINST INHALED
 CC MICROORGANISMS. BINDS STRONGLY MALTOSE RESIDUES AND TO A LESSER
 CC EXTENT OTHER ALPHA-GLUCOSYL MOETIES. IT COULD PARTICIPATE IN THE
 CC EXTRACELLULAR REORGANIZATION OR TURNOVER OF PULMONARY SURFACTANT.
 CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 4 SET OF HOMOTRIMERS.
 CC -1- SUBCELLULAR LOCATION: Extracellular.
 CC -1- PTM: The N-terminus is blocked.
 CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
 CC PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
 CC CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
 CC HYDROPHOBIC PROTEINS (SP-B AND SP-C).
 CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: L05485; AAB59450.1; -
 CC EMBL: L05483; AAB59450.1; JOINED.
 CC EMBL: L05484; AAB59450.1; JOINED.
 CC EMBL: X65018; CAA46152.1; -
 CC PIR: A45225; A45225.
 CC PIR: S18382; S18382.
 CC PDB: 1B08; 29-NOV-99.
 CC Genew; HGNC:10803; SFTPD.
 CC MIM: 178635; -
 CC InterPro: IPR000087; Collagen.
 CC InterPro: IPR001304; Lectin_C.
 CC Pfam: PF00059; Lectin_C; 1.
 CC Pfam: PF01391; Collagen; 4.
 CC SMART: SM00034; CLECT; 1.
 CC PROSITE: PS00615; C-TYPE LECTIN_1; 1.
 CC PROSITE: PS50041; C-TYPE LECTIN_2; 1.
 KW Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
 KW Signal; Lectin; Collagen; Repeat; Collid coll; 3D-structure.

```

FT SIGNAL 1 20 BY SIMILARITY.
FT CHAIN 21 375 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT 46 222 D.
FT DOMAIN 223 252 COLLAGEN-LIKE.
FT 279 375 COILED COIL (POTENTIAL).
FT DISULFID 281 373 C-TYPE LECTIN (SHORT FORM).
FT CARBOHYD 351 365 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT MOD.RES 90 90 HYDROXYLATION (BY SIMILARITY).
FT 78 78 HYDROXYLATION (BY SIMILARITY).
FT MOD.RES 96 96 HYDROXYLATION (BY SIMILARITY).
FT MOD.RES 99 99 HYDROXYLATION (BY SIMILARITY).
FT MOD.RES 171 171 HYDROXYLATION (BY SIMILARITY).
FT MOD.RES 177 177 HYDROXYLATION (BY SIMILARITY).
FT CONFLICT 31 31 M -> T (IN REF. 2).
FT CONFLICT 59 59 P -> F (IN REF. 3).
FT CONFLICT 122 122 A -> P (IN REF. 2).
FT CONFLICT 180 180 T -> A (IN REF. 2).
FT CONFLICT 206 206 E -> P (IN REF. 3).
FT CONFLICT 374 374 E -> EH (IN REF. 3).
SQ SEQUENCE 375 AA; 37702 MW; 2986B2699FC01A6A CRC64;

Query Match 4.5%; Score 11; DB 1; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.074;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDGR 55
Db 46 GLPGRDGRDGR 56

RESULT 7
CA21_RANCA STANDARD; PRT: 1355 AA.
AC 042350;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Collagen alpha 2(I) chain precursor.
GN COL1A2.
OS Rana catesbeiana (Bull. frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Neobatrachia; Ranoidae; Ranidae; Rana.
OX NCBI_TaxID=8400;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tail;
RX MEDLINE=97417499; PubMed=9272872;
RA Asahina K., Oofusa K., Obara M., Yoshizato K.;
RT Cloning and characterization of the full length cDNA encoding alpha2
RT type I collagen of bullfrog Rana catesbeiana.;
RL Gene 1947263-289(1997).
CC -1- FUNCTION: TYPE I COLLAGEN IS A MEMBER OF GROUP I COLLAGEN
CC (FIBRILLAR FORMING COLLAGEN).
CC -1- SUBUNIT: TRIMERS OF ONE ALPHA 2(I) AND TWO ALPHA 1(I) CHAINS.
CC -1- TISSUE SPECIFICITY: FORMS THE FIBRILS OF TENDON, LIGAMENTS AND
CC BONES. IN BONES THE FIBRILS ARE MINERALIZED WITH CALCIUM
CC HYDROXYAPATITE.
CC -1- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D88764; BAA22380.1;
CC InterPro: IPR000087; Collagen.
CC InterPro: IPR000885; Fib_collagen_C.

```

```

DR Pfam: PF01391; Collagen; 18.
DR Pfam: PF01410; COLFI; 1.
DR ProDom: PD000007; Collagen; 3.
DR ProDom: PD002078; Fib_collagen_C; 1.
DR SMART: SM00038; COLFI; 1.
KW Extracellular matrix; Connective tissue; Repeat; Hydroxylation;
KW Glycoprotein; Collagen; Signal.
FT SIGNAL 1 22 POTENTIAL.
FT PROPEP 23 82 AMINO-TERMINAL PROPEPTIDE (POTENTIAL).
FT CHAIN 83 1093 COLLAGEN ALPHA 2(I) CHAIN.
FT PROPEP 1094 1355 CARBOXYL-TERMINAL PROPEPTIDE
FT (BY SIMILARITY).
FT CARBOHYD 1206 1206 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 1256 1256 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 1355 AA; 127643 MW; CB793AD5D6F41D2A CRC64;

Query Match 4.5%; Score 11; DB 1; Length 1355;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 83 RGEAGPAGPTG 93
Db 705 RGEAGPAGPTG 715

RESULT 8
CA34_HUMAN STANDARD; PRT: 1670 AA.
AC 001955; Q9B0T2;
DT 01-OCT-1996 (Rel. 34, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Collagen alpha 3(IV) chain precursor (Goodpasture antigen).
GN COL4A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=94364994; PubMed=8083201;
RA Maruyama M., Leinonen A., Mochizuki T., Tryggvason K., Reiders S.T.;
RT "Complete primary structure of the human alpha 3(IV) collagen chain.
RT Coexpression of the alpha 3(IV) and alpha 4(IV) collagen chains in
RT human tissues.";
RL J. Biol. Chem. 269:23013-23017(1994).
RN [2]
RP REVISIONS.
RA Leinonen A.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.; VARIANTS AS E-297; R-407; R-640; R-1167; E-1207;
RP 0-1215; S-1277; T-1330; E-1334; E-1347 AND C-1661, AND VARIANTS R-43;
RP E-162; Y-326; H-408; R-451; L-574; E-1269 AND P-1474.
RX MEDLINE=21064696; PubMed=1134255;
RA Heider L., Aronidel C., Forestier L., Cohen-solal L., Mollet G.,
RA Gutierrez B., Stavrou C., Gubler M.C., Antignac C.;
RT "Structure of the human type IV collagen gene COL4A3 and mutations in
RT autosomal Alport syndrome.";
RL J. Am. Soc. Nephrol. 12:97-106(2001).
RN [4]
RP SEQUENCE OF 1386-1670 FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=93015826; PubMed=1400291;
RA Quiñones S., Bernal D., Garcia-Sogo M., Elena S.F., Saus J.;
RT "Exon/intron structure of the human alpha 3(IV) gene encompassing the
RT Goodpasture antigen (alpha 3(IV)NC1). Identification of a potentially
RT antigenic region at the triple helix/NC1 domain junction.";
RL J. Biol. Chem. 267:19780-19784(1992).
RN [5]
RP SEQUENCE OF 1453-1670 FROM N.A.
RX MEDLINE=91353570; PubMed=1882840;
RA Morrison K.E., Maruyama M., Yang-Feng T.L., Reiders S.T.;

```

RT *Sequence and localization of a partial cDNA encoding the human alpha
RT 3 chain of type IV collagen.";
RL Am. J. Hum. Genet. 49:545-554(1991).
RN [6]
RP SEQUENCE OF 1331-1670 FROM N.A.
RC TISSUE-Kidney;
RX MEDLINE-92147878; PubMed-1737849;
RA Turner N., Mason P.J., Brown R., Fox M., Povey S., Rees A.,
RA Pusey C.D.;
RT *Molecular cloning of the human Goodpasture antigen demonstrates it
RT to be the alpha 3 chain of type IV collagen.";
RL J. Clin. Invest. 89:552-601(1992).
RN [7]
RP SEQUENCE OF 1644-1670 FROM N.A.
RC TISSUE-Kidney;
RA Ding J.;
RL Submitted (JAN-1993) to the EMBL/GenBank/DBJ databases.
RN [8]
RP SEQUENCE OF 1439-1670, AND ALTERNATIVE SPLICING.
RC TISSUE-Kidney;
RX MEDLINE-94124597; PubMed-8294492;
RA Feng L., Xia Y., Wilson C.B.;
RT *Alternative splicing of the NCI domain of the human alpha 3(IV)
RT collagen gene. Differential expression of mRNA transcripts that
RT predict three protein variants with distinct carboxyl regions.";
RL J. Biol. Chem. 269:2342-2348(1994).
RN [9]
RP SEQUENCE OF 1-29 FROM N.A.
RX MEDLINE-98196854; PubMed-9537506;
RA Momota R., Sugimoto M., Ohashi T., Kigasawa K., Yoshioke H.,
RA Nishimura Y.;
RT *Two genes, COL4A3 and COL4A4 coding for the human alpha3(IV) and
RT alpha4(IV) collagen chains are arranged head-to-head on chromosome
RT 2q36.";
RL FEBS Lett. 424:11-16(1998).
RN [10]
RP ALTERNATIVE SPLICING.
RX MEDLINE-93280184; PubMed-8505332;
RA Bernal D., Gullones S., Saus J.;
RT *The human mRNA encoding the Goodpasture antigen is alternatively
RT spliced";
RL J. Biol. Chem. 268:12090-12094(1993).
RN [11]
RP VARIANT PRO-1474.
RX MEDLINE-95078827; PubMed-7987301;
RA Lemmink H.H., Mochizuki T., van den Heuvel L.P.W.J., Schroeder C.H.,
RA Bartelentos A., Monnens L.A.H., van Oost B.A., Brunner H.G.,
RA Reeders S.T., Smeets H.J.M.;
RT *Mutations in the type IV collagen alpha 3 (COL4A3) gene in autosomal
RT recessive Alport syndrome.";
RL Hum. Mol. Genet. 3:1269-1273(1994).
RN [12]
RP FUNCTION: TYPE IV COLLAGEN IS THE MAJOR STRUCTURAL COMPONENT OF
RN GLOMERULAR BASEMENT MEMBRANES (GBM), FORMING A 'CHICKEN-WIRE'
RN MESHWORK TOGETHER WITH LAMININS, PROTEOGLYCANS AND ENACTIN/
RN NIDOGEN.
RN [13]
RP SUBUNIT: THERE ARE SIX TYPE IV COLLAGEN ISOFORMS, ALPHA 1(IV)-
RN ALPHA 6(IV), EACH OF WHICH CAN FORM A TRIPLE HELIX STRUCTURE
RN WITH 2 OTHER CHAINS TO GENERATE TYPE IV COLLAGEN NETWORK.
RN [14]
RP SUBCELLULAR LOCATION: CELL SURFACE (POTENTIAL).
RN [15]
RP ALTERNATIVE PRODUCTS: AT LEAST 3 ISOFORMS; 1 (SHOWN HERE), 2/V AND
RN 3/LS; ARE PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER IN THEIR
RN C-TERMINAL NCI DOMAINS.
RN [16]
RP TISSUE SPECIFICITY: ALPHA 3 AND ALPHA 4 TYPE IV COLLAGENS ARE
RN COLOCALIZED AND PRESENT ONLY IN BASEMENT MEMBRANES OF KIDNEY, EYE,
RN COCHLEA, LUNG AND BRAIN.
RN [17]
RP DOMAIN: ALPHA CHAINS OF TYPE IV COLLAGEN HAVE A NONCOLLAGENOUS
RN DOMAIN (NCI) AT THEIR C-TERMINUS, FREQUENT INTERRUPTIONS OF THE
RN G-X-Y REPEATS IN THE LONG CENTRAL TRIPLE-HELICAL DOMAIN (WHICH MAY
RN CAUSE FLEXIBILITY IN THE TRIPLE HELIX), AND A SHORT N-TERMINAL
RN TRIPLE-HELICAL 7S DOMAIN.
RN [18]
RP PTM: PROLINS ARE AT THE THIRD POSITION OF THE TRIPLEPEIDE REPEATING
RN UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
RN [19]
RP PTM: THE ALTERNATIVE SPLICED FORM V CONTAINS AN ADDITIONAL

CC N-LINKED GLYCOSYLATION SITE.
CC -1- PTM: TYPE IV COLLAGENS CONTAIN NUMEROUS CYSTEINE RESIDUES WHICH
CC ARE INVOLVED IN INTER- AND INTRAMOLECULAR DISULFIDE BONDING. 12 OF
CC THESE, LOCATED IN THE NCI DOMAIN, ARE CONSERVED IN ALL KNOWN TYPE
CC IV COLLAGENS.
CC -1- PTM: Phosphorylated by the Goodpasture antigen-binding protein.
CC -1- DISEASE: ANTIBODIES AGAINST THE NCI DOMAIN OF ALPHA3(IV) MEDIATE
CC THE HUMAN AUTOIMMUNE DISEASE, GOODPASTURE SYNDROME, WHICH IS
CC CHARACTERIZED BY HEMATURIA AND PULMONARY HEMORRHAGE.
CC -1- DISEASE: DEFECTS IN COL4A3 ARE ASSOCIATED WITH THE TYPE I
CC AUTOSOMAL RECESSIVE FORM OF ALPORT SYNDROME, AN HEREDITARY
CC GLOMERULONEPHROPATHY CHARACTERIZED BY PROGRESSIVE RENAL FAILURE,
CC HEMATURIA AND DEAFNESS. THE RECESSIVE FORM OCCURS EQUALLY BETWEEN
CC MALES AND FEMALES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-slb.ch/announce/>
CC or send an email to license@isb-slb.ch).
CC -----
CC EMBL; X80031; CAA56335.1; -
CC EMBL; AJ288487; CAC36101.1; JOINED.
CC EMBL; AJ288488; CAC36101.1; JOINED.
CC EMBL; AJ288489; CAC36101.1; JOINED.
CC EMBL; AJ288490; CAC36101.1; JOINED.
CC EMBL; AJ288491; CAC36101.1; JOINED.
CC EMBL; AJ288492; CAC36101.1; JOINED.
CC EMBL; AJ288493; CAC36101.1; JOINED.
CC EMBL; AJ288494; CAC36101.1; JOINED.
CC EMBL; AJ288495; CAC36101.1; JOINED.
CC EMBL; AJ288496; CAC36101.1; JOINED.
CC EMBL; AJ288497; CAC36101.1; JOINED.
CC EMBL; AJ288498; CAC36101.1; JOINED.
CC EMBL; AJ288499; CAC36101.1; JOINED.
CC EMBL; AJ288500; CAC36101.1; JOINED.
CC EMBL; AJ288501; CAC36101.1; JOINED.
CC EMBL; AJ288502; CAC36101.1; JOINED.
CC EMBL; AJ288503; CAC36101.1; JOINED.
CC EMBL; AJ288504; CAC36101.1; JOINED.
CC EMBL; AJ288505; CAC36101.1; JOINED.
CC EMBL; AJ288506; CAC36101.1; JOINED.
CC EMBL; AJ288507; CAC36101.1; JOINED.
CC EMBL; AJ288508; CAC36101.1; JOINED.
CC EMBL; AJ288509; CAC36101.1; JOINED.
CC EMBL; AJ288510; CAC36101.1; JOINED.
CC EMBL; AJ288511; CAC36101.1; JOINED.
CC EMBL; AJ288512; CAC36101.1; JOINED.
CC EMBL; AJ288513; CAC36101.1; JOINED.
CC EMBL; AJ288514; CAC36101.1; JOINED.
CC EMBL; AJ288515; CAC36101.1; JOINED.
CC EMBL; AJ288516; CAC36101.1; JOINED.
CC EMBL; AJ288517; CAC36101.1; JOINED.
CC EMBL; AJ288518; CAC36101.1; JOINED.
CC EMBL; AJ288519; CAC36101.1; JOINED.
CC EMBL; AJ288520; CAC36101.1; JOINED.
CC EMBL; AJ288521; CAC36101.1; JOINED.
CC EMBL; AJ288522; CAC36101.1; JOINED.
CC EMBL; AJ288523; CAC36101.1; JOINED.
CC EMBL; AJ288524; CAC36101.1; JOINED.
CC EMBL; AJ288525; CAC36101.1; JOINED.
CC EMBL; AJ288526; CAC36101.1; JOINED.
CC EMBL; AJ288527; CAC36101.1; JOINED.
CC EMBL; AJ288528; CAC36101.1; JOINED.
CC EMBL; AJ288529; CAC36101.1; JOINED.
CC EMBL; AJ288530; CAC36101.1; JOINED.
CC EMBL; AJ288531; CAC36101.1; JOINED.
CC EMBL; AJ288532; CAC36101.1; JOINED.
CC EMBL; AJ288533; CAC36101.1; JOINED.
CC EMBL; AJ288534; CAC36101.1; JOINED.
CC EMBL; AJ288535; CAC36101.1; JOINED.

DR EMBL: AJ288536; CAC36101.1; JOINED.
 DR EMBL: AJ288537; CAC36101.1; JOINED.
 DR EMBL: AJ288538; CAC36101.1; JOINED.
 DR EMBL: M92993; AAA21610.1; -.
 DR EMBL: S55790; AAB19637.1; -.
 DR EMBL: M81379; AAA51556.1; -.
 DR EMBL: L08650; AAA52044.1; -.
 DR EMBL: U02519; AAA18942.1; -.

Query Match 4.5%; Score 11; DB 1; Length 1670;
 Best Local Similarity 100.0%; Pred. No. 0.24;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 71 PGLPGRGDPG 81
 DB 1300 PGLPGRGDPG 1310

RESULT 9
 CA17_HUMAN STANDARD; PRT; 2944 AA.
 AC 002388; 014054; 016507;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Collagen alpha 1(VII) chain precursor (Long-chain collagen) (LC collagen).
 GN COL7A1.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-94327588; PubMed-8051117;
 RA Christiano A.M., Greenspan D.S., Lee S., Uitto J.;
 RT Cloning of human type VII collagen. Complete primary sequence of the alpha 1(VII) chain and identification of intergenic polymorphisms.
 RT J. Biol. Chem. 269:20256-20262(1994).
 RN [2]
 RP SEQUENCE OF 128-1493 FROM N.A., AND PARTIAL SEQUENCE.
 RX MEDLINE-93336437; PubMed-1307247;
 RA Christiano A.M., Rosenbaum L.M., Chung-Honet L.C., Parente M.G., Woodley D.T., Pan T.C., Zhang R.Z., Chu M.-L., Burgeson R.E., Uitto J.;
 RT The large non-collagenous domain (NC-1) of type VII collagen is amino-terminal and chimeric. Homology to cartilage matrix protein, the type III domains of fibronectin and the A domains of von Willebrand factor.
 RT Hum. Mol. Genet. 1:475-481(1992).
 RN [3]
 RP SEQUENCE OF 815-1439 FROM N.A.
 RX MEDLINE-93107742; PubMed-1871109;
 RA Parente M.G., Chung L.C., Rymaenen J., Woodley D.T., Wynn K.W., Bauer E.A., Mettel M.-G., Chu M.-L., Uitto J.;
 RT Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.
 RT Proc. Natl. Acad. Sci. U.S.A. 88:6931-6935(1991).
 RN [4]
 RP SEQUENCE OF 369-1255 FROM N.A.
 RX MEDLINE-93107742; PubMed-1469284;
 RA Gammon W.R., Abernethy M.L., Padilla K.M., Pitsayanh P.S., Cook M.E., Wright J., Brigaman R.A., Hunt S.W. III;
 RT Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion proteins involved in tissue-specific organization of extracellular matrix.
 RT J. Invest. Dermatol. 99:691-696(1992).
 RN [5]
 RP SEQUENCE OF 340-675 FROM N.A.
 RX TISSUE-Keratinocytes; PubMed-1567409;
 RA Tanaka T., Takahashi K., Furukawa F., Imamura S.;
 RT Molecular cloning and characterization of type VII collagen cDNA.

RL Biochem. Biophys. Res. Commun. 183:958-963(1992).
 RN [6]
 RP SEQUENCE OF 2395-2944 FROM N.A.
 RX MEDLINE-93271985; PubMed-8499916;
 RA Greenspan D.S.;
 RT The carboxyl-terminal half of type VII collagen, including the non-collagenous NC-2 domain and intron/exon organization of the corresponding region of the COL7A1 gene.
 RT Hum. Mol. Genet. 2:273-278(1993).
 RN [7]
 RP SEQUENCE OF 1-87 FROM N.A.
 RX TISSUE-Placenta;
 RC MEDLINE-94375010; PubMed-8088784;
 RA Christiano A.M., Hoffman G.G., Chung-Honet L.C., Lee S., Cheng W., Uitto J., Greenspan D.S.;
 RT Structural organization of the human type VII collagen gene (COL7A1), composed of more exons than any previously characterized gene.
 RT Genomics 21:169-179(1994).
 RN [8]
 RP REVIEW ON DEB VARIANTS.
 RX MEDLINE-98041696; PubMed-9375948;
 RA Jaervikallio A., Pulkkinen L., Uitto J.;
 RT Molecular basis of dystrophic epidermolysis bullosa: mutations in the type VII collagen gene (COL7A1).
 RT Hum. Mutat. 10:338-347(1997).
 RN [9]
 RP VARIANT RDEB LYS-2798.
 RX MEDLINE-93291877; PubMed-8513326;
 RA Christiano A.M., Greenspan D.S., Hoffman G.G., Zhang X., Tamai Y., Lin A.N., Dietz H.C., Hovnanian A., Uitto J.;
 RT A missense mutation in type VII collagen in two affected siblings with recessive dystrophic epidermolysis bullosa.
 RT Nat. Genet. 4:62-66(1993).
 RN [10]
 RP VARIANT DDEB SER-2040.
 RX MEDLINE-94224777; PubMed-8170945;
 RA Christiano A.M., Rymaenen M., Uitto J.;
 RT Dominant dystrophic epidermolysis bullosa: identification of a Gly-->Ser substitution in the triple-helical domain of type VII collagen.
 RT Proc. Natl. Acad. Sci. U.S.A. 91:3549-3553(1994).
 RN [11]
 RP VARIANT PER-DDEB CYS-2623.
 RX MEDLINE-96081220; PubMed-8541842;
 RA Christiano A.M., Lee J.Y.-Y., Chen W.J., Laforgia S., Uitto J.;
 RT Peribulbar epidermolysis bullosa: genetic linkage to COL7A1 and identification of a glycine-to-cysteine substitution in the triple-helical domain of type VII collagen.
 RT Hum. Mol. Genet. 4:1579-1583(1995).
 RN [12]
 RP VARIANT DDEB ARG-2043.
 RX MEDLINE-95164985; PubMed-7861014;
 RA Christiano A.M., Morriconi A., Paradisi M., Angelo C., Mazzanti C., Cavallieri R., Uitto J.;
 RT A glycine-to-arginine substitution in the triple-helical domain of type VII collagen in a family with dominant dystrophic epidermolysis bullosa.
 RT J. Invest. Dermatol. 104:438-440(1995).
 RN [13]
 RP VARIANTS RDEB AND DDEB.
 RX MEDLINE-96220218; PubMed-8644729;
 RA Christiano A.M., McGrath J.A., Tan K.C., Uitto J.;
 RT Glycine substitutions in the triple-helical region of type VII collagen result in a spectrum of dystrophic epidermolysis bullosa phenotypes and patterns of inheritance.
 RT Am. J. Hum. Genet. 58:671-681(1996).
 RN [14]
 RP VARIANT RDEB ARG-2575.
 RX MEDLINE-96154068; PubMed-8592061;
 RA Shimizu H., McGrath J.A., Christiano A.M., Nishikawa T., Uitto J.;
 RT Molecular basis of recessive dystrophic epidermolysis bullosa: genotype/phenotype correlation in a case of moderate clinical severity.

RL J. Invest. Dermatol. 106:119-124(1996).
 RN [15]
 RP VARIANT RDEB ARC-1782.
 RX MEDLINE-96183562; PubMed-8618018;
 RA Christiano A.M., McGrath J.A., Uitto J.;
 RT "Influence of the second COL7A1 mutation in determining the
 RT phenotypic severity of recessive dystrophic epidermolysis bullosa."
 RL J. Invest. Dermatol. 106:766-770(1996).
 RN [16]
 RP VARIANT RDEB ASP-2073.
 RX MEDLINE-96310789; PubMed-8757758;
 RA Dunnill M.G.S., McGrath J.A., Richards A.J., Christiano A.M.,
 RA Uitto J., Pope F.M., Rady R.A.J.;
 RT "Clinicopathological correlations of compound heterozygous COL7A1
 RT mutations in recessive dystrophic epidermolysis bullosa."
 RL J. Invest. Dermatol. 107:171-177(1996).
 RN [17]
 RP VARIANTS RDEB W-1982; G-2008; A-2025; E-2049; G-2063; W-2063 AND
 RP R-2575.
 RX MEDLINE-97465605; PubMed-9326325;
 RA Hovanian A., Rochat A., Bodemer C., Petit E., Rivers C.A., Prost C.,
 RA Fraitag S., Christiano A.M., Uitto J., Lathrop M., Barrandon Y.,
 RA de Prost Y.;
 RT "Characterization of 18 new mutations in COL7A1 in recessive
 RT dystrophic epidermolysis bullosa provides evidence for distinct
 RT molecular mechanisms underlying defective anchoring fibril
 RT formation."
 RL Am. J. Hum. Genet. 61:599-610(1997).
 RN [18]
 RP VARIANT RDEB ARC-1652.
 RX MEDLINE-98106792; PubMed-9444387;
 RA Cserhalmi-Friedman P.B., Karpatt S., Horvath A., Christiano A.M.;
 RT "Identification of a glycine substitution and a splice site mutation
 RT in the type VII collagen gene in a proband with milds recessive
 RT dystrophic epidermolysis bullosa."
 RL Arch. Dermatol. Res. 289:640-645(1997).
 RN [19]
 RP VARIANTS DEB ARC-2009 AND ARC-2043.
 RX MEDLINE-97358588; PubMed-9215684;
 RA Weinberg J.-O., Hammami-Hausall N., Nijssen O., Anton-Lamprecht I.,
 RA Naylor S.L., Kerbacher K., Zimmermann M., Krajci P.,
 RA Gedde-Dahl T., Bruckner-Tuderman L.;
 RT "Modulation of disease severity of dystrophic epidermolysis bullosa by
 RT a splice site mutation in combination with a missense mutation in the
 RT COL7A1 gene."
 RL Hum. Mol. Genet. 6:1125-1135(1997).
 RN [20]
 RP VARIANTS DEB ASP-1519; ASP-2006; GLU-2015 AND ARC-2034.
 RX MEDLINE-98334662; PubMed-9668111;
 RA Hammami-Hausall N., Schumann H., Raghunath M., Kilgus O., Luethi U.,
 RA Lager T., Bruckner-Tuderman L.;
 RT "Some, but not all, glycine substitution mutations in COL7A1 result in
 RT intracellular accumulation of collagen VII, loss of anchoring
 RT fibrils, and skin blistering."
 RL J. Biol. Chem. 273:19228-19234(1998).
 RN [21]
 RP VARIANTS DEB CYS-2008; ARC-2207 AND SER-2775.
 RX MEDLINE-98410969; PubMed-9740253;
 RA Kon A., Pulkkinen L., Ishida-Yamamoto A., Hashimoto I., Uitto J.;
 RT "Novel COL7A1 mutations in dystrophic forms of epidermolysis
 RT bullosa."
 RL J. Invest. Dermatol. 111:534-537(1998).
 RN [22]
 RP VARIANT RDEB ARC-1347.
 RX MEDLINE-99019477; PubMed-9804332;
 RA Terracina M., Postoraro P., Schubert M., Sonego G., Atzori F.,
 RA Zambruno G., Bruckner-Tuderman L., Castiglia D.;
 RT "Compound heterozygosity for a recessive glycine substitution and a
 RT splice site mutation in the COL7A1 gene causes an unusually mild form
 RT of localized recessive dystrophic epidermolysis bullosa."
 RL J. Invest. Dermatol. 111:744-750(1998).
 RN [23]
 RP VARIANTS DEB TRP-2034; VAL-2040; ARG-2043; ARG-2064 AND ASP-2713.

RX MEDLINE-99072663; PubMed-9856843;
 Query Match 4.5%; Score 11; DB 1; Length 2944;
 Best Local Similarity 100.0%; Pred. No. 0.38;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 71 PGLPGRGDPG 81
 Db 1328 PGLPGRGDPG 1338
 RESULT 10
 CAC2_HAECO
 ID CAC2_HAECO STANDARD; PRT; 210 AA.
 AC P16252;
 DT 01-AUG-1990 (Rel. 15, Created)
 DT 01-AUG-1990 (Rel. 15, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Cuticle collagen 2C (Fragment).
 GN 2C.
 OS Haemonchus contortus (Barber pole worm).
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
 OC Trichostrongyloidea; Haemonchidae; Haemonchinae; Haemonchus.
 OX NCBI_TaxID=6289;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-90136718; PubMed-2615789;
 RA Shamsky L.M., Pratt D., Bolsven R.J., Cox G.N.;
 RT "Cuticle collagen genes of Haemonchus contortus and Caenorhabditis
 RT elegans are highly conserved."
 RL Mol. Biochem. Parasitol. 37:73-86(1989).
 CC -1- FUNCTION: NEMATODE CUTICLES ARE COMPOSED LARGELY OF COLLAGEN-LIKE
 CC PROTEINS. THE CUTICLE FUNCTIONS BOTH AS AN EXOSKELETON AND AS A
 CC BARRIER TO PROTECT THE WORM FROM ITS ENVIRONMENT.
 CC -1- MISCELLANEOUS: THIS PROTEIN SHOWS 4 POTENTIAL TRIPLE-HELICAL
 CC REGIONS, WHICH CONTAIN GLYCINE AS EVERY THIRD AMINO ACID.
 CC -1- MISCELLANEOUS: IN ALL NEMATODE CUTICLE COLLAGENS, THE POLYPEPTIDE
 CC CHAINS ARE COMPLEXED WITHIN THE CUTICLE BY DISULFIDE BONDS AND
 CC OTHER TYPES OF COVALENT CROSS-LINKS.
 CC -1- SIMILARITY: BELONGS TO THE CUTICULAR COLLAGEN FAMILY.
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL J04670; AAA29172.1; -
 DR Interpro: IPR000087; Collagen.
 DR Pfam: PF01391; Collagen; 3.
 KM Cuticle; Connective tissue; Repeat; Multigene family; Collagen.
 FT NON TER 1 1
 SO SEQUENCE 210 AA; 19562 MW; E15FA9A2BD3D74B CRC64;
 Query Match 4.1%; Score 10; DB 1; Length 210;
 Best Local Similarity 100.0%; Pred. No. 0.31;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 57 GARGPGRGK 66
 Db 151 GARGPGRGK 160
 RESULT 11
 PSPA_CAVPO
 ID PSPA_CAVPO STANDARD; PRT; 247 AA.
 AC P50403;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A)

DE (PSP).
GN SFTPA1 OR SFTPA OR SFTPI.
OS Cavia porcellus (Guinea pig).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Hystriognath; Cavidae; Cavia.
OX NCBI_TaxID=10141;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-Hartley; TISSUE-Lung;
RX MEDLINE=98018900; PubMed=9357868;
RA Yuan H.T., Gowan S., Kelly F.J., Bingle C.D.;
RT Cloning of guinea pig surfactant protein A defines a distinct
cellular distribution pattern within the lung.*
RL Am. J. Physiol. 273:L900-L906(1997).
CC -1- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSPA BINDS TO SURFACTANT
PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE
AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS
ESSENTIAL FOR NORMAL RESPIRATION.
CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC -1- SIMILARITY: CONTRAINS A COLLAGENOUS DOMAIN.
CC -1- SIMILARITY: CONTRAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC EMBL: U04869; AAB82952.1; -
DR HSSP: P22897; IEGG.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C; 1.
DR Pfam: PF01391; Collagen; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE: PS00411; C-TYPE_LECTIN_2; 1.
KW Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW Signal; Lectin; Collagen; Repeat.
FT SIGNAL 1 19 POTENTIAL.
FT CHAIN 20 247 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT FT
FT DOMAIN 27 99 A.
FT FT
FT DOMAIN 152 245 COLLAGEN-LIKE.
FT FT
FT DISULFID 154 245 C-TYPE LECTIN (SHORT FORM).
FT FT
FT DISULFID 223 237 BY SIMILARITY.
FT FT
FT CARBOHYD 20 20 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT FT
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 247 AA: 26104 MW: 2866270EFC932 CRC64;
Query Match 4.1%; Score 10; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DE Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A)
GN SFTPA1 OR SFTPA OR SFTPI.
OS Oryctolagus cuniculus (Rabbit).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN (1)
RP SEQUENCE FROM N.A.
RC MEDLINE=88139348; PubMed=2830270;
RX Boggsam V., Qing K., Mendelson C.R.;
RA "The major apoprotein of rabbit pulmonary surfactant. Elucidation of
primary sequence and cyclic AMP and developmental regulation.*"
RL J. Biol. Chem. 263:2933-2947(1988).
CC [2]
RN SEQUENCE FROM N.A.
RC STRAIN-New Zealand white; TISSUE-Liver;
RX MEDLINE=92312742; PubMed=1616051;
RA Chen O., Boggsam V., Mendelson C.R.;
RT "Rabbit lung surfactant protein A gene: Identification of a lung-
specific DNAase I hypersensitive site.*"
RL Am. J. Physiol. 263:L662-L671(1992).
CC -1- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSPA BINDS TO SURFACTANT
PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE
AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS
ESSENTIAL FOR NORMAL RESPIRATION.
CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10%
PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS,
CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL
HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC -1- SIMILARITY: CONTRAINS A COLLAGENOUS DOMAIN.
CC -1- SIMILARITY: CONTRAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).
CC EMBL: J03542; AAA1465.1; -
DR EMBL: L19387; AAA31468.1; -
DR PIR: A29311; LNRPS.
DR HSSP: P22897; IEGG.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF00059; Lectin_C; 1.
DR Pfam: PF01391; Collagen; 2.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.
DR PROSITE: PS00411; C-TYPE_LECTIN_2; 1.
KW Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation;
KW Signal; Lectin; Collagen; Repeat.
FT SIGNAL 1 15 POTENTIAL.
FT CHAIN 16 247 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT FT
FT DOMAIN 27 99 A.
FT FT
FT DOMAIN 152 247 COLLAGEN-LIKE.
FT FT
FT DISULFID 154 245 C-TYPE LECTIN (SHORT FORM).
FT FT
FT DISULFID 223 237 BY SIMILARITY.
FT FT
FT CARBOHYD 206 206 N-LINKED (GLCNAC. . .) (PROBABLE).
FT FT
FT VARIANT 12 12 S -> P.
FT CONFLICT 57 60 GPMG -> APWA (IN REF. 2).
SQ SEQUENCE 247 AA: 26071 MW: 289634054C8C8CB4 CRC64;
Query Match 4.1%; Score 10; DB 1; Length 247;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 39 GLPGRDGRDG 48

RESULT 13
PSPA_CANFA STANDARD: PRT: 248 AA.
AC P06908: 01-JAN-1988 (Rel. 06, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A) (PSP-A)
GN SFTPA1 OR SFTPA OR SFTP1
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
NCBI_TaxID=9615;
[1]
RP SEQUENCE FROM N.A. AND SIGNAL SEQUENCE CLEAVAGE SITE.
RX MEDLINE=66016705; PubMed=3863100;
RA Benson B., Hawgood S., Schilling J., Clements J., Damm D., Cordell B., White R.T.;
RT "Structure of canine pulmonary surfactant apoprotein: cDNA and complete amino acid sequence."
RL Proc. Natl. Acad. Sci. U.S.A. 82:6379-6383(1985).
[2]
RN C-TYPE LECTIN DOMAIN.
RP MEDLINE=87115834; PubMed=3808053;
RX Palty L.;
RT "Is lung surfactant protein a lectin-collagen hybrid?";
RL Nature 325:490-490(1987).
CC -1- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSPA BINDS TO SURFACTANT PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS ESSENTIAL FOR NORMAL RESPIRATION.
CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10% PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS, CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M11769; AAA30887.1; -
CC PIR: A25296; LNDGFS.
CC HSSP: P22897; 1EGG.
CC InterPro: IPR000087; Collagen.
CC InterPro: IPR001304; LECTIN_C.
CC Pfam: PF00059; LECTIN_C; 1.
CC Pfam: PF01391; Collagen; 2.
CC SMART: SM00034; CLECT; 1.
CC PROSITE: PS00615; C-TYPE LECTIN_1; 1.
CC PROSITE: PS00615; C-TYPE LECTIN_2; 1.
CC Glycoprotein; Calcium; Surface film; Gaseous exchange; Hydroxylation; Signal; Lectin; Collagen; Repeat.
KW CHAIN 1 17
FT FT 18 248 PULMONARY SURFACTANT-ASSOCIATED PROTEIN
FT FT 1 17 A.
FT FT 28 100 COLLAGEN-LIKE.
FT FT 153 248 C-TYPE LECTIN (SHORT FORM).
FT FT 155 246 BY SIMILARITY.
FT FT 224 238 BY SIMILARITY.
FT FT 20 20 N-LINKED (GLCNAC...) (POTENTIAL).
FT CARBOHYD

FT CARBOHYD 207 207 N-LINKED (GLCNAC...) (PROBABLE).
SQ SEQUENCE 248 AA; 26268 MW; 340FE95D4E2502C0 CRC64;
Query Match 4.18; Score 10; DB 1; Length 248;
Best Local Similarity 100.0%; Pred. No. 0.35;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 45 GLPGRDGRDG 54
Db 40 GLPGRDGRDG 49
RESULT 14
PSPA_HUMAN STANDARD: PRT: 248 AA.
AC P07714:
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Pulmonary surfactant-associated protein A precursor (SP-A) (PSP-A) (PSP-A)
DE (PSP-A) (Alveolar proteinosis protein) (35 kDa pulmonary surfactant-associated protein).
GN SFTPA1 OR SFTPA OR SFTP1 OR PSPA.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NCBI_TaxID=9606;
[1]
RN C-TYPE LECTIN DOMAIN.
RP SEQUENCE FROM N.A.
RX MEDLINE=66250832; PubMed=3755136;
RA Erioz J., Steinbrink R., Jacobs K., Phelps D., Kriz R., Recny M., Sultzman L., Jones S., Tausch H.W., Frank H.A., Fritsch E.F.;
RT "Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfactant-associated protein."
RL J. Biol. Chem. 261:9029-9033(1986).
[2]
RN C-TYPE LECTIN DOMAIN.
RP SEQUENCE FROM N.A.
RX MEDLINE=86014366; PubMed=2995821;
RA White R.T., Damm D., Miller J., Spratt K., Schilling J., Hawgood S., Benson B., Cordell B.;
RT "Isolation and characterization of the human pulmonary surfactant apoprotein gene."
RL Nature 317:361-363(1985).
[3]
RN C-TYPE LECTIN DOMAIN.
RP SEQUENCE FROM N.A.
RX MEDLINE=92198680; PubMed=1372511;
RA Katyal S.L., Singh G., Locker J.L.;
RT "Characterization of a second human pulmonary surfactant-associated protein SP-A gene."
RL Am. J. Respir. Cell Mol. Biol. 6:446-452(1992).
CC -1- FUNCTION: IN PRESENCE OF CALCIUM IONS, PSPA BINDS TO SURFACTANT PHOSPHOLIPIDS AND CONTRIBUTES TO LOWER THE SURFACE TENSION AT THE AIR-LIQUID INTERFACE IN THE ALVEOLI OF THE MAMMALIAN LUNG AND IS ESSENTIAL FOR NORMAL RESPIRATION.
CC -1- SUBUNIT: OLIGOMERIC COMPLEX OF 6 SET OF HOMOTRIMERS.
CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- MISCELLANEOUS: PULMONARY SURFACTANT CONSISTS OF 90% LIPID AND 10% PROTEIN. THERE ARE 4 SURFACTANT ASSOCIATED PROTEIN: 2 COLLAGENOUS, CARBOHYDRATE-BINDING GLYCOPROTEINS (SP-A AND SP-D) AND 2 SMALL HYDROPHOBIC PROTEINS (SP-B AND SP-C).
CC -1- SIMILARITY: CONTAINS A COLLAGENOUS DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC -----
CC EMBL: M13686; AAA60211.1; -
CC EMBL: K03475; AAA36520.1; -

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 14:48:42 ; Search time 39 Seconds
(Without alignments)
598.991 Million cell updates/sec

Title: US-09-943-851A-42

Perfect score: 1325

Sequence: 1 MRPLVLLLLGLAAGSPPLD.....DSTFGFLVYSDMHSSPVFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Capext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08
Maximum Match 1008

Listing first 45 summaries

Database : PIR.73:*

1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1207	91.1	219	2	T14782
2	424	32.0	244	2	JC4708
3	418.5	31.6	680	1	CGHUI1D
4	416.5	31.4	674	2	S23297
5	412	31.1	246	2	S29328
6	411.5	31.1	674	2	S13301
7	408	30.8	680	2	S31216
8	399	30.1	245	1	C1HUOC
9	374	28.2	744	1	A34246
10	370	27.9	744	1	S23298
11	363.5	27.4	743	1	S23779
12	363	27.3	635	3	A57131
13	362	27.3	253	1	C1HUOB
14	360.5	27.2	253	3	S49158
15	359	27.1	744	2	S15435
16	351	26.5	253	2	I49560
17	316	23.8	423	2	A55797
18	314	23.7	245	1	C1HUOA
19	314	23.7	245	1	S19018
20	280.5	21.2	215	2	B48150
21	277.5	20.9	215	2	C48150
22	261.5	19.7	196	2	A48150
23	238.5	18.0	992	2	T08772
24	230.5	17.4	1049	1	CG8075
25	227.5	17.4	248	2	E15921
26	225.5	17.0	248	1	LNHUP1
27	225	17.0	335	2	T32248
28	224	16.9	636	2	S41067
29	223	16.8	1464	2	S59856

30	222.5	16.8	1758	2	T29350	hypothetical prote
31	222.5	16.8	1759	2	T29351	collagen alpha 2(I
32	222	16.8	360	2	T37285	collagen dpy-2 - C
33	221.5	16.7	170	2	B57131	collagen alpha 2(V
34	221	16.7	886	2	I50694	collagen alpha 1(I
35	221	16.7	1019	1	A32856	collagen alpha 1(I
36	220	16.6	248	1	LNDGFS	pulmonary surfacta
37	220	16.6	341	2	T16296	hypothetical prote
38	220	16.6	380	2	T28888	cuticle collagen d
39	219.5	16.6	247	1	LNRBPS	pulmonary surfacta
40	219.5	16.6	381	2	T27806	hypothetical prote
41	219.5	16.6	673	1	CG806C	collagen alpha 1(I
42	219.5	16.6	1466	1	CGHUTL	collagen alpha 1(I
43	219	16.5	671	1	CGRTIS	collagen alpha 1(I
44	218.5	16.5	1453	2	S21626	collagen alpha 1(I
45	217	16.4	283	2	T29980	hypothetical prote

ALIGNMENTS

```

RESULT 1
T14782
hypothetical protein DKFZp586B0621.1 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 02-Sep-2000
C:Accession: T14782
R:Ottewaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, August 1999
A:Reference number: Z18184
A:Accession: T14782
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-219 <OT>
A:Cross-references: EMBL:AL110261
A:Experimental source: adult uterus; clone DKFZp586B0621
C:Genetics:
A>Note: DKFZp586B0621.1
C:Superfamily: complement C1q carboxyl-terminal homology

Query Match          91.1%; Score 1207; DB 2; Length 219;
Best Local Similarity 100.0%; Pred. No. 1.3e-77;
Matches 219; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 25 PSLCPGHCGLPRTPGHHSGSLPGRDGRGARGARGEGGRPGLPGRGDPGPRG 84
DB 1 PSLCPGHPGLPCTPGHHSGSLPGRDGRGARGARGEGGRPGLPGRGDPGPRG 60
QY 85 EAGPAGPTGPGAGCGSVPPSAFSAKRSERVPSPDAPLPFDRLVNEQGHYDAVTGKFT 144
DB 61 EAGPAGPTGPGAGCGSVPPSAFSAKRSERVPSPDAPLPFDRLVNEQGHYDAVTGKFT 120
QY 145 CQVPGVYFAVHATYVRSALQFDLVKNGESIASFFQFGGWPCKPASLSGAMVRLPEPDQ 204
DB 121 CQVPGVYFAVHATYVRSALQFDLVKNGESIASFFQFGGWPCKPASLSGAMVRLPEPDQ 180
QY 205 VVWVGVDYIGITVSIKTDSTFGFLVYSDMHSSPVFA 243
DB 181 VVWVGVDYIGITVSIKTDSTFGFLVYSDMHSSPVFA 219

RESULT 2
JC4708
gelatin-binding 28k protein precursor - human
N:Alternate names: adipose specific collagen-like factor
C:Species: Homo sapiens (man)
C:Date: 10-May-1996 #sequence_revision 19-Jul-1996 #text_change 20-Sep-1999
C:Accession: JC4708; Jc4944
R:Maeda, K.; Okubo, K.; Shimomura, I.; Funahashi, T.; Matsuzawa, Y.; Matsubara, K.
Biochem. Biophys. Res. Commun. 221, 286-289, 1996
A:Title: cDNA cloning and expression of a novel adipose specific collagen-like factor
A:Reference number: JC4708; MUID:96224171; PMID:8619847
A:Accession: JC4708

```

A:Molecule type: mRNA
A:Residues: 1-244 <NAE>
A:CROSS-references: DDBJ:D45371; NID:g871886; PIDD:BAA08227.1; PID:g871887
A:Experimental source: adipose tissue
R:Nakano, Y.; Tohe, T.; Choi-Miyata, N.H.; Mazda, T.; Tomita, M.
J. Biochem. 120, 803-812, 1996
A>Title: Isolation and characterization of GBP2, a novel gelatin-binding protein purified from rat brown adipocytes
A:Accession number: JG4944; MUID:97103474; PMID:8947845
A:Molecule type: Protein
A:Residues: 19-38;93-100;101-112;135-149;173-178 <NAK>
C:Comment: This protein is an endogenous factor that binds with a collagen-like domain.
C:Genetics:
A:Gene: apml
C:Superfamily: unassigned collagens; complement C1q carboxyl-terminal homology
C:Keywords: adipose tissue; glycoprotein; hydroxyproline
F:1-18/Domin: signal sequence #status predicted <SIG>
F:19-244/Product: gelatin-binding 28kDa protein #status experimental <MAT>
F:42-107/Region: collagen-like
F:114-241/Domin: complement C1q carboxyl-terminal homology <CIQ>
F:95/Modified site: 4-hydroxyproline (Pro) #status experimental
F:230/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 32.0%; Score 424; DB 2; Length 244;
Best Local Similarly 40.5%; Pred. No. 7,7e-23;
Matches 106; Conservative 26; Mismatches 72; Indels 58; Gaps 10;

DY 6 VLLLLGLAAGSPPLDDNKKIPSLCPGH-----PG-----LPGTGGHGSQ 44
| | | | | | | | | | | | | | | | | | : | | | |
DY 7 VLLLLAL-----PQHDETTTQGPGVLLPLPKACKICGMMAIGTGHCHN 50
| | | | | | | | | | | | | | | | | | : | | | |
DY 45 GLPRGRDGRGDGAPGAPGEKEGEGRPGLPRGPD-----PGRGEAGPATGPA 95
| | | | | | | | | | | | | | | | | | : | | | |
DY 51 GAPRGDRD-----GPGEKEGKDGRLIGPKKDGETGVGAEGRGFGIGRKCEP 104
| | | | | | | | | | | | | | | | | | : | | | |
DY 96 GEGCVPRPSAFSAKSRSRRVPESDAPLPPRVLVNQGHDAVTCKPTQCVGVVVFAV 155
| | | | | | | | | | | | | | | | | | : | | | |
DY 105 GEGGYVYSRASV--GLEYYVTP--NPPIREFKITLYNDQNTDSGTGFHCNIPGLTYFAY 162
| | | | | | | | | | | | | | | | | | : | | | |
DY 156 HATYVRASLFDFLVKNSES-IASFQFFGCGMPKPASISGCAMVBLEBEDQWVGV- 213
| | | | | | | | | | | | | | | | | | : | | | |
DY 163 HITVYMMDVKVSLERKKOKAMFTYDQYQENNVDQA--SGSVLHLLEVDQWVLQYGE 220
| | | | | | | | | | | | | | | | | | : | | | |
DY 214 YIGIVASIKTDSTFSGFLVSD 235
| | | | | | | | | | | | | | | | | | : | | | |
DY 221 RNLGYADNDNDSTFTGFLHYD 242
| | | | | | | | | | | | | | | | | | : | | | |

RESULT 3
CGRID
collagen alpha 1(x) chain precursor - human
N:Alternate names: procollagen alpha 1(x) chain
C:Species: Homo sapiens (man)
C>Date: 22-Nov-1993 #sequence revision 03-Nov-1995 #text change 22-Jun-1999
C:Accession: S26396; S30086; S15826; S18249; A43901; I51870; S21856
R:Reichenberger, E.; Beller, F.; LuValle, P.; Olsen, B.R.; von der Mark, K.; Bertling, W.
FEBS Lett. 311, 305-310, 1992
A>Title: Genomic organization and full-length cDNA sequence of human collagen X.
A:Reference number: S26396; MUID:93012005; PMID:1397333
A:Accession: S26396
A:Molecule type: DNA
A:Residues: 1-680 <RET>
A:CROSS-references: EMBL:X60952; EMBL:X72578; EMBL:X72579; EMBL:X72580; GB:S47714; GB:S47715
R:Apte, S.S.
submitted to the EMBL Data Library, March 1992
A:Reference number: S30085
A:Accession: S30086
A:Molecule type: DNA
A:Residues: 'RTTFGWCMVCLL', 52-680 <APT>
A:CROSS-references: EMBL:X65120; NID:g233129
A>Note: The initial difference is probably due to translation of an intronic sequence
R:Apte, S.; Mattei, M.G.; Olsen, B.R.
FEBS Lett. 282, 393-396, 1991

A>Title: Cloning of human alpha-1(X) collagen DNA and localization of the COL10A1 gene
A:Reference number: S15826; MUID:91243838; PMID:2037056
A:Accession: S15826
A:Molecule type: DNA
A:Residues: 561-647, 'G', 649-666 <AP2>
A:Cross-references: EMBL:X58879; NID:930013; PIDN:CAA41686.1; PID:930014
R:Thomas, J.T.; Cresswell, C.J.; Rash, B.; Nicolai, H.; Jones, T.; Solomon, E.; Grant
Biochem. J. 280, 617-623, 1991
A>Title: The human collagen X gene. Complete primary translated sequence and chromoso
A:Reference number: S18249; MUID:92109659; PMID:1764025
A:Accession: S18249
A:Molecule type: DNA
A:Residues: 1-26, 'T', 28-680 <THO>
A:Cross-references: EMBL:X60382; NID:930094; PIDN:CAA42933.1; PID:930095
A>Note: the sequence from Fig. 3 is inconsistent with that from Fig. 2 in having 532-
R:Reichenberger, E.; Aigner, T.; von der Mark, K.; Stoss, H.; Bertling, W.
Dev. Biol. 148, 562-572, 1991
A>Title: In situ hybridization studies on the expression of type X collagen in fetal
A:Reference number: A43901; MUID:92077285; PMID:1743401
A:Accession: A43901
A:Molecule type: mRNA
A:Residues: 547-656 <RE2>
A:Cross-references: GB:M474050; GB:D57494; NID:9339884; PIDN:AA61221.1; PID:9553796
A>Note: sequence extracted from NCBI backbone (NCBIIN:69012, NCBIAP:69014)
R:Wallis, G.A.; Rash, B.; Sweetman, W.A.; Thomas, J.T.; Super, M.; Evans, G.; Grant,
Am. J. Hum. Genet. 54, 169-178, 1994
A>Title: Amino acid substitutions of conserved residues in the carboxyl-terminal doma
pe Schmid.
A:Reference number: I51870; MUID:94136476; PMID:8304336
A:Accession: I51870
A>Status: translated from GB/EMBL/DBDB
A:Molecule type: mRNA
A:Residues: 520-597, 'D', 599-680 <MAL>
A:Cross-references: GB:568531; NID:9545180; PIDN:AAC60615.1; PID:9545181
A>Note: mutant sequence from patient with metaphyseal chondrodysplasia type Schmid
A>Note: a second mutant sequence with 614-Pro is also described
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
ed and subsequently O-glycosylated.
C:Genetics:
A:Gene: GDB:COL10A1
A:Cross-references: GDB:128635; OMIM:120110
A:Map position: 6q21-6q22
A:Introns: 52/1
A>Note: a defect in this gene may cause Schmid metaphyseal chondrodysplasia
C:Complex: type X collagen may be a homotrimer
C:Function:
A:Description: structural component of extracellular fibrous polymer specifically and
be important for skeletogenesis
C:Superfamily: collagen alpha 1(VIII) chain; complement C1q carboxyl-terminal homolog
C:Keywords: coiled coil; extracellular matrix; glycoprotein; homotrimer; hydroxyllysine
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-680/Product: collagen alpha 1(X) chain #status predicted <MAT>
F:19-56/Domain: amino-terminal nonhelical #status predicted <NC2>
F:57-519/Region: interrupted helical
F:552-680/Domain: amino-terminal nonhelical #status predicted <NC1>
F:553-679/Domain: complement C1q carboxyl-terminal homology <C1Q>
F:617/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match	31.6%	Score 418.5	DB 1	Length 680
Best Local Similarity	34.9%	Pred. No. 5,4e-22		
Matches	98	Conservative 34	Mismatches 82	Indels 67
		Gaps 6		
DB	15	GSPLDNDKIPSLCGRHGLGTPGHHGSGGLGGRDGRDGRDGPARGAGEGCGRGLP	74	
DB	401	GNPGLPGRPKGDPGVGGPGLGPRVGPARGAGMGCHMGCHNDEAGPRGAGIPIGRGPIRGPGPIR	460	
QY	75	-----GPRDGPGRGEA-----GPAGPTGPAECES	99	
DB	461	GFPGSKGDDPGSPGPPGAGIATKGLNGTPGPPGPPGPRGSHSGEGLGPGPGPPGQAV	520	
QY	100	VP-----FRSAFSAKRSRSNRVPPSDADLPDRVLVNE	132	
DB	521	MPGCFIRAGORPSLSTGPLVANSAGVGTGMPVASTATVLLSKAV--PALGTPIPEPKILLYNR	578	

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 14:40:55 ; Search time 83 Seconds
(without alignments)
390.119 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 1325
Sequence: 1 MRLVLVLLGLGLAASPPLD.....DSFGSGFLVSDWHSVPFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues
Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: A.GeneSeq-101002:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
24: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1325	100.0	243	AAV06481	Human tumour-associ
2	1325	100.0	243	AAV17827	Human PRO344 prote
3	1325	100.0	243	AAW97984	Human adipocyte-sp
4	1325	100.0	243	AA833451	Human PRO344 prote
5	1325	100.0	243	AAV17168	Human PRO344 prote
6	1325	100.0	243	AAV93668	Human acid sequenc
7	1325	100.0	243	AA801318	Human PRO344 polyp
8	1325	100.0	243	AAU12352	Human PRO344 polyp
9	1325	100.0	243	AA85815	Human TANGO 253 SE
10	1325	100.0	243	AA849593	Human adipocyte co

11	1325	100.0	243	AA849599	Human adipocyte co
12	1323	99.8	243	AA85891	Human secreted pro
13	1321	99.7	243	AA85888	Human secreted pro
14	1321	99.7	243	AA85889	Human secreted pro
15	1321	99.7	243	AA85890	Human secreted pro
16	1271	95.9	243	AAV76040	Rat skin cell prote
17	1271	95.9	243	AA855979	Skin cell protein,
18	1271	95.9	243	AA872179	Rat protein isolat
19	1258	94.9	243	AA858820	Murine TANGO 253 S
20	1256	94.8	228	AA85816	Murine mature TANG
21	1256	94.8	243	AA85899	Murine secreted pr
22	1254	94.6	243	AA85897	Murine secreted pr
23	1254	94.6	243	AA85898	Murine secreted pr
24	1243.5	93.8	242	AA85896	Murine secreted pr
25	1200	90.6	228	AA858821	Murine mature TANG
26	1183	89.3	220	ABG12724	Novel human diagn
27	1048	79.1	201	AAO21663	Human secreted pro
28	675	50.9	128	AA85819	Human TANGO 253 C1
29	650	49.1	128	AA85824	Murine TANGO 253 C
30	478	36.1	151	ABG12723	Novel human diagn
31	449.5	33.9	225	AB880583	Human sbg1033026C1
32	442.5	33.4	243	AB880582	Human sbg1033026C1
33	433.5	32.7	247	AA809107	Murine adipocyte c
34	433.5	32.7	247	AA805528	Mouse OB33 protein
35	433.5	32.7	247	AB808232	Mouse acrp30 prote
36	425	32.1	244	AA809108	Human adipocyte co
37	424.5	32.0	247	AA805527	Mouse OB33 protein
38	424.5	32.0	247	AB808221	Mouse adipocyte prote
39	424	32.0	244	AAV21807	Adipose most abund
40	424	32.0	244	AA830233	Human adipocyte co
41	424	32.0	244	AAV71035	Human APM1 (Adipos
42	424	32.0	244	AA805529	Human OB33 protein
43	424	32.0	244	AA85828	Human OB33 protein
44	424	32.0	244	AA849592	Human ACRP30 prote
45	424	32.0	244	AA849598	Human ACRP30 prote

ALIGNMENTS

RESULT 1	AAV06481	AAV06481 standard; Protein: 243 AA.
ID	AAV06481	
XX	AAV06481	
AC	AAV06481	
XX	27-SEP-1999 (first entry)	
XX	Human tumour-associated protein PRO344.	
DE	PRO344; UNQ303; cancer; tumour; diagnosis; therapy; human.	
XX		
XX	Homo sapiens.	
OS		
XX		
FH	key	Location/Qualifiers
FT	Peptide	1..15
FT	/note="signal peptide"	
FT	Protein	16..243
FT	/note="mature protein"	
FT	Modified-site	68..215
FT	/note="N-myristoylated"	
FT	Modified-site	216..243
FT	/note="N-myristoylated"	
XX		
XX	MO9935170-A2.	
PN	15-JUL-1999.	
XX		
PD	05-JAN-1999;	99MO-US00106.
XX		
PF	20-NOV-1998;	98US-0109304.
XX	05-JAN-1998;	98US-0070440.
PR	29-APR-1998;	98US-0083500.

PR 22-MAY-1998; 98US-0086414.
 PR 10-JUN-1998; 98US-0088742.
 PR 10-NOV-1998; 98US-0107783.
 XX
 PA (GETH) GENENTECH INC.
 PI Botstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA;
 PI Roy MA, Wood WI;
 XX
 DR WPI: 1999-430385/36.
 DR N-PSDB: AAX87258.
 XX
 PT Antibody against proteins expressed in neoplastic cells, useful for
 PT tumor diagnosis and treatment
 XX
 PS Example 1, Fig 10; 162pp; English.
 CC This sequence represents human PRO344 (UNQ303), a protein encoded
 CC by the novel cDNA clone DNA40592 (see AAX87258). Amplification of
 CC DNA40592 was observed in primary lung tumours and in primary colon
 CC tumours, suggesting a significant role in tumour formation and
 CC growth. Antagonists (e.g. antibodies) directed to PRO344 may have
 CC use in cancer therapy. The invention identifies 14 genes (see
 CC AAX87254-67) that are amplified in the genome of tumour cells. Such
 CC amplification is expected to be associated with overexpression of
 CC the gene product and to contribute to tumorigenesis. The encoded
 CC proteins (see AAY06477-90) may be useful targets for the diagnosis
 CC and/or treatment (including prevention) of certain cancers, and may
 CC act as predictors of the prognosis of tumour treatment. Antibodies
 CC that bind the proteins are claimed and used in claimed cancer
 CC diagnostic kits.
 CC
 XX
 SQ Sequence 243 AA;
 Query Match 100.0%; Score 1325; DB 20; Length 243;
 Best Local Similarity 100.0%; Pred. NO. 5.3e-103;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLLLGLAAGSPRLDNRKIPSLCPGHPGLPCTPGHNSGGLPGRDGRDGRGAPG 60
 DB 1 MRPLVLLLLGLAAGSPRLDNRKIPSLCPGHPGLPCTPGHNSGGLPGRDGRDGRGAPG 60
 QY 61 APGKGGGGRGRLPGRGDPGRGEGAPGPTGACGSCVPPRSATSAKSESRRVPPSD 120
 DB 61 APGKGGGGRGRLPGRGDPGRGEGAPGPTGACGSCVPPRSATSAKSESRRVPPSD 120
 QY 121 APLPFDRLVNEOGHYDAVTGKFTCOVPGYYFAVHATYVRASLQFDLVNKGESIASFFQ 180
 DB 121 APLPFDRLVNEOGHYDAVTGKFTCOVPGYYFAVHATYVRASLQFDLVNKGESIASFFQ 180
 QY 181 FFGGMPKPAASLGGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSFSGFLVYSDMHSSP 240
 DB 181 FFGGMPKPAASLGGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSFSGFLVYSDMHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243
 XX
 RESULT 2
 ID AAY17827 standard; Protein; 243 AA.
 XX
 AC AAY17827;
 XX
 DT 12-AUG-1999 (first entry)
 XX
 DE Human PRO344 protein sequence.
 XX
 KW Human; PRO protein; tumour necrosis factor family; TNF; cytokine;
 KW secreted protein; transmembrane protein; inflammation disorder.
 XX
 OS Homo sapiens.

XX
 PN WO928462-A2.
 XX
 PD 10-JUN-1999.
 XX
 PF 01-DEC-1998; 98MO-US25108.
 XX
 PR 25-FEB-1998; 98US-0075945.
 PR 03-DEC-1997; 97US-0067411.
 PR 11-DEC-1997; 97US-0069278.
 PR 11-DEC-1997; 97US-0069334.
 PR 11-DEC-1997; 97US-0069335.
 PR 12-DEC-1997; 97US-0069425.
 PR 16-DEC-1997; 97US-0069694.
 PR 16-DEC-1997; 97US-0069696.
 PR 16-DEC-1997; 97US-0069702.
 PR 17-DEC-1997; 97US-0069870.
 PR 18-DEC-1997; 97US-0069873.
 PR 05-JAN-1998; 98US-0070440.
 PR 09-FEB-1998; 98US-0074086.
 PR 09-FEB-1998; 98US-0074092.
 XX
 PA (GETH) GENENTECH INC.
 XX
 PI Baker KP, Chen J, Goddard A, Gurney AL, Wood WI;
 PI Yuan J;
 XX
 DR WPI: 1999-371118/31.
 DR N-PSDB: AAX80052.
 XX
 PT Nucleic acids encoding PRO secreted and transmembrane proteins
 PS Claim 12; Fig 21; 123pp; English.
 CC The present invention describes nucleic acids encoding PRO secreted and
 CC transmembrane proteins used therapeutically. The PRO proteins have
 CC cytosolic, anti-inflammatory, anti-proliferative and immunosuppressive
 CC activity. The proteins and polynucleotides can be used in therapy,
 CC identification of homologues, raising antibodies and design of probes
 CC and primers. They can be used in a range of diseases related to proteins
 CC that they have homology with, e.g. a PRO protein having homology to
 CC complement proteins may be used in inflammatory responses.
 CC
 XX
 SQ Sequence 243 AA;
 Query Match 100.0%; Score 1325; DB 20; Length 243;
 Best Local Similarity 100.0%; Pred. NO. 5.3e-103;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLLLGLAAGSPRLDNRKIPSLCPGHPGLPCTPGHNSGGLPGRDGRDGRGAPG 60
 DB 1 MRPLVLLLLGLAAGSPRLDNRKIPSLCPGHPGLPCTPGHNSGGLPGRDGRDGRGAPG 60
 QY 61 APGKGGGGRGRLPGRGDPGRGEGAPGPTGACGSCVPPRSATSAKSESRRVPPSD 120
 DB 61 APGKGGGGRGRLPGRGDPGRGEGAPGPTGACGSCVPPRSATSAKSESRRVPPSD 120
 QY 121 APLPFDRLVNEOGHYDAVTGKFTCOVPGYYFAVHATYVRASLQFDLVNKGESIASFFQ 180
 DB 121 APLPFDRLVNEOGHYDAVTGKFTCOVPGYYFAVHATYVRASLQFDLVNKGESIASFFQ 180
 QY 181 FFGGMPKPAASLGGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSFSGFLVYSDMHSSP 240
 DB 181 FFGGMPKPAASLGGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSFSGFLVYSDMHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243
 XX
 RESULT 3
 AAN97984

ID AAM97984 standard; Protein; 243 AA.
XX AAM97984;
XX
XX
XX 21-JUN-1999 (first entry)
XX
XX Human adipocyte-specific protein zsig39.
XX
XX Adipocyte-specific protein; zsig39; human; fatty acid metabolism;
XX energy balance; nutrition; antimicrobial.
XX
XX Homo sapiens.
FH Key Location/Qualifiers
FT Peptide 1..15 "signal peptide, alternatively the signal
FT /note- "peptide comprises residues 1..18"
FT Protein 16..243
FT /note- "mature protein, alternatively the mature
FT protein comprises residues 19..243
FT (specifically claimed in Claim 4)"
FT Domain 30..96
FT /note- "collagen-like domain"
FT Domain 98..243
FT /note- "globular domain"
FT Region 105..109
FT /note- "beta strand"
FT Region 128..130
FT /note- "beta strand"
FT Region 136..139
FT /note- "beta strand"
FT Region 143..146
FT /note- "beta strand"
FT Region 164..171
FT /note- "beta strand"
FT Region 176..182
FT /note- "beta strand"
FT Region 187..200
FT /note- "beta strand"
FT Region 204..210
FT /note- "beta strand"
FT Region 226..231
FT /note- "beta strand"
FT Domain 111..135
FT /note- "receptor binding domain"
FT Domain 170..174
FT /note- "receptor binding domain"
XX
XX MO9910492-A1.
XX
XX 04-MAR-1999.
XX
XX 26-AUG-1998; 98MO-US17724.
XX
XX 26-AUG-1997; 97US-0056983.
XX
XX (ZYMO) ZYMOGENETICS INC.
XX
XX Humes JM, Sheppard PO;
XX
XX MPI: 1999-204665/17.
XX
XX N-PSDB: AAX24684.
XX
XX
XX zsig39 protein - used to modulate fatty acid metabolism
XX
XX Claim 1; 111-112; 132pp; English.
XX
XX This polypeptide comprises human adipocyte-specific protein zsig39,
XX a protein that modulates free fatty acid metabolism. zsig39 is a
XX member of a family of proteins having a globular domain and a
XX collagen-like domain capable of dimerisation or oligomerisation.
XX zsig39 polypeptides were initially identified by querying an EST
XX database for secretory signal sequences characterised by an upstream

CC methionine start site, a hydrophobic region of approximately 13
CC amino acids and a cleavage site. A single EST sequence was
CC discovered, and the novel polypeptide encoded by the full-length
CC cDNA allowed the identification of a homologue relationship with
CC adipocyte complement related protein Acyr30 and adipocyte secreted
CC protein apM1. A full-length clone (see AAX24684) was obtained from a
CC lung tissue library. Expression vectors, cultured cells and a
CC method of producing zsig39 polypeptide are claimed, as well as
CC zsig39 polypeptides having N- or C-terminal affinity tags, toxins,
CC radionuclides, enzymes or fluorophores, fusion proteins
CC including zsig39 polypeptides, an antibody that specifically
CC binds to an epitope of zsig39, and a method for modulating free
CC fatty acid metabolism by administering a zsig39 polypeptide. The
CC zsig39 polypeptide may also be used in organ preservation, for
CC cryopreservation, for surgical pretreatment to prevent injury due
CC to ischaemia and/or inflammation, and as an antimicrobial agent,
CC promoting lysis or phagocytosis of infectious agents.
XX
XX
SQ Sequence 243 AA:
Query Match 100.0%; Score 1325; DB 20; Length 243;
Best Local Similarity 100.0%; Pred. No. 5.3e-103;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPLVLLLLGLAAGSPPLDNRKIPSLCPGHGGLPCTPGHSGSLPGHGRGDRGARG 60
DB 1 MRPLVLLLLGLAAGSPPLDNRKIPSLCPGHGGLPCTPGHSGSLPGHGRGDRGARG 60
QY 61 APEGKGGGRGPGLPGRGDPGRGEAGPAGPGCEGCVPPRPAFAKRSRRPPSD 120
DB 61 APEGKGGGRGPGLPGRGDPGRGEAGPAGPGCEGCVPPRPAFAKRSRRPPSD 120
QY 121 APLPEPRVLYNQGHDATGKFTCCVPGVYFAVATYRRASLDLKKNESTASFPQ 180
DB 121 APLPEPRVLYNQGHDATGKFTCCVPGVYFAVATYRRASLDLKKNESTASFPQ 180
QY 181 FFGWPKPASLSGCAVRLPEPDQVWVQGVGDYIGIVASIKTDSFFSGFLVSDMHSSP 240
DB 181 FFGWPKPASLSGCAVRLPEPDQVWVQGVGDYIGIVASIKTDSFFSGFLVSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243
RESULT 4
AAB33461
ID AAB33461 standard; Protein; 243 AA.
XX
XX AAB33461;
XX
XX 29-JAN-2001 (first entry)
XX
XX Human PRO344 protein UNQ303 SEQ ID NO:241.
XX
XX Human; immune related disease; diagnosis; antinflammatory; cardiac;
XX dermatological; antiarthritic; antirheumatic; immunosuppressive;
XX haemostatic; antithyroid; antidiabetic; neuroprotective;
XX antinaeemic; hepatotropic; virucide; antipsoriatic; antiallergic;
XX antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;
XX osteoarthritis; spondyloarthritis; systemic sclerosis; sarcoidosis;
XX idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
XX systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
XX autoimmune thrombocytopenia; immune-mediated renal disease;
XX demyelinating disease; hepatobiliary disease; Whipple's disease;
XX inflammatory bowel disease; gluten-sensitive enteropathy;
XX autoimmune disease; immune-mediated skin disease; allergic disease;
XX immunological disease; transplantation associated disease;
XX graft rejection; graft-versus-host-disease.
XX
XX Homo sapiens.
XX
XX MO200053758-A2.
PN

XX 14-SEP-2000.
 XX 02-MAR-2000; 2000MO-US05841.
 XX 08-MAR-1999; 99MO-US05028.
 PR 10-MAR-1999; 99US-0123618.
 PR 12-MAR-1999; 99US-0123957.
 PR 23-MAR-1999; 99US-0125775.
 PR 12-APR-1999; 99US-0128849.
 PR 20-APR-1999; 99MO-US08615.
 PR 28-APR-1999; 99US-0131445.
 PR 04-MAY-1999; 99US-0132371.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99MO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145688.
 PR 28-JUL-1999; 99US-0146222.
 PR 01-SEP-1999; 99MO-US20111.
 PR 08-SEP-1999; 99MO-US20594.
 PR 13-SEP-1999; 99MO-US20944.
 PR 15-SEP-1999; 99MO-US21090.
 PR 15-SEP-1999; 99MO-US21547.
 PR 05-OCT-1999; 99MO-US23089.
 PR 29-OCT-1999; 99US-0162506.
 PR 29-NOV-1999; 99MO-US28214.
 PR 30-NOV-1999; 99MO-US28313.
 PR 30-NOV-1999; 99MO-US28409.
 PR 01-DEC-1999; 99MO-US28301.
 PR 01-DEC-1999; 99MO-US28634.
 PR 02-DEC-1999; 99MO-US28551.
 PR 02-DEC-1999; 99MO-US28564.
 PR 16-DEC-1999; 99MO-US30095.
 PR 20-DEC-1999; 99MO-US30999.
 PR 30-DEC-1999; 99MO-US31274.
 PR 05-JAN-2000; 2000MO-US00219.
 PR 06-JAN-2000; 2000MO-US00277.
 PR 11-FEB-2000; 2000MO-US00376.
 PR 11-FEB-2000; 2000MO-US03565.
 PR 18-FEB-2000; 2000MO-US04341.
 PR 22-FEB-2000; 2000MO-US04342.
 PR 22-FEB-2000; 2000MO-US04414.
 (GETH) GENENTECH INC.
 XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;
 PI Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;
 PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M;
 XX N-PSDB; AAC58626.
 DR MPI: 2000-572271/53.
 XX Sixty four PRO polypeptides, useful in the diagnosis and treatment of
 PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
 XX
 PS Claim 33: Fig 96; 309pp; English.
 XX
 CC The present invention describes sixty four human PRO proteins which can
 CC be used in the treatment of immune related diseases. The human PRO
 CC proteins, anti-PRO antibodies, agonists and antagonists are useful for
 CC treating and diagnosing immune related disorders. The disorders are
 CC selected from systemic lupus erythematosus, rheumatoid arthritis,
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthritis,
 CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
 CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,
 CC immune-mediated renal disease, demyelinating diseases of the central
 CC and peripheral nervous systems, hepatobiliary diseases, inflammatory
 CC bowel disease, gluten-sensitive enteropathy and Whipple's disease,
 CC autoimmune or immune-mediated skin diseases, allergic diseases,

CC immunological diseases of the lung, and transplantation associated
 CC diseases including graft rejection and graft-versus-host-disease.
 CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and
 CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.
 XX
 XX Sequence 243 AA;
 SQ
 Query Match 100.0%; Score 1325; DB 21; Length 243;
 Best Local Similarity 100.0%; Pred. No. 5, 3e-103;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLLLGLAASPPDDNKRIPSLCPGHPGLPPTPGHHGSQGLGRDGRDRCAGP 60
 DB 1 MRPLVLLLLGLAASPPDDNKRIPSLCPGHPGLPPTPGHHGSQGLGRDGRDRCAGP 60
 QY 61 APECKEGBGRPGLPGRGRDPRGEGAGPAGTPTGACSVPPRSATSAKRSERVPPSPD 120
 DB 61 APECKEGBGRPGLPGRGRDPRGEGAGPAGTPTGACSVPPRSATSAKRSERVPPSPD 120
 QY 121 APLEPDRVLVNEGGHTDAVYGRFTCOVPGYTRAVHATYRASLQEDLVNKGESIASFFQ 180
 DB 121 APLEPDRVLVNEGGHTDAVYGRFTCOVPGYTRAVHATYRASLQEDLVNKGESIASFFQ 180
 QY 181 FFGWPKRPASLSCGAVRLEPEDQVWVGVGDYIGIVASIKTDSFFSGFLVSDMHSSP 240
 DB 181 FFGWPKRPASLSCGAVRLEPEDQVWVGVGDYIGIVASIKTDSFFSGFLVSDMHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243
 RESULT 5
 AAY71468
 ID AAY71468 standard; Protein; 243 AA.
 XX
 AC AAY71468;
 XX
 DT 08-NOV-2000 (first entry)
 XX
 DE Human PRO344 protein.
 XX
 KW PRO344; DNA40592-1242; human; ATCC No: 209492; antiproliferative;
 KW neoplastic cell growth inhibitor; cytostatic; treatment: cancer; tumour;
 KW breast; prostate; colon; lung; renal; ovarian; central nervous system;
 KW CNS; leukemia; melanoma; Expressed Sequence Tag; EST; secreted protein;
 KW extracellular domain; ECD.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..15
 FT Modified-site /label= signal_peptide
 FT /note= "N-myristoylation site"
 FT Protein 16..243
 FT /label= Mature_PRO344_protein
 FT Modified-site 68..74
 FT /note= "N-myristoylation site"
 FT Binding-site 77..80
 FT /note= "Cell attachment sequence"
 FT Modified-site 216..222
 FT /note= "N-myristoylation site"
 PN MO200032778-A2.
 XX
 PD 08-JUN-2000;
 XX
 PD 30-NOV-1999; 99MO-US28409.
 XX
 PR 01-DEC-1998; 99MO-US25108.

PR 16-DEC-1998; 98US-0112850.
 PR 22-DEC-1998; 98US-0113296.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.

PA (GETH) GENENTECH INC.

PI Chen J, Goddard A, Gurney AL, Hillan K, Napier M, Wood WI;

XX MPI: 2000-412325/35.

DR N-PSDB; AAD01241.

XX New composition useful for inhibiting neoplastic cell growth and for
 PT treating cancers, comprises PRO655, PRO344 or PRO364 polypeptide or
 PT their antagonists

XX Claim 31: Fig 6; 108pp; English.

XX The present sequence is the human PRO344 protein, encoded by the cDNA
 CC clone, designated as DNA40592-1242. It is isolated from human foetal
 CC lung tissue, cDNA library, identified using probes based on a consensus
 CC sequence DNA34398, derived from secreted protein extracellular domain
 CC (ECD) expressed sequence tag (EST). This clone is assigned ATCC deposit
 CC No: 209492. PRO344 functions as a neoplastic cell growth inhibitor and
 CC is used for treating tumours, using an effective amount of PRO655, PRO364
 CC and PRO344. This composition is especially useful for treatment of human
 CC cancers such as breast, prostate, colon, lung, renal, ovarian and CNS,
 CC leukemia and melanoma.

XX Sequence 243 AA:

Query Match 100.0%; Score 1325; DB 21; Length 243;

Best Local Similarity 100.0%; Pred. No. 5.3e-103; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTPGHNSOGLPGHGRDGRDGRGARG 60
 DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTPGHNSOGLPGHGRDGRDGRGARG 60
 QY 61 ARGEKGGGRPLPGRCGDPGRGEGPAGPTGPAECVPRPSAFSARSSSRVPPSPD 120
 DB 61 ARGEKGGGRPLPGRCGDPGRGEGPAGPTGPAECVPRPSAFSARSSSRVPPSPD 120
 QY 121 APLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATYRASLOFDLVKNGESIASFQ 180
 DB 121 APLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATYRASLOFDLVKNGESIASFQ 180
 QY 181 FFGWPKPASISGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLYSDMHSSP 240
 DB 181 FFGWPKPASISGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLYSDMHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243

RESULT 6

AA93688 ID AA93688 standard; Protein; 243 AA.

XX AA93688;

XX 03-OCT-2000 (first entry)

XX Amino acid sequence of novel polypeptide PRO344.

XX PRO301; PRO327; PRO1265; PRO344; PRO347; PRO357;
 KW PRO315; PRO1017; PRO1112; PRO509; PRO882; tumour cell;
 KM tumorigenesis; cancer; neoplastic cell growth; cell proliferation.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Peptide 1..15
 FT /note= "signal peptide"
 FT Modified-site 11..17
 FT /note= "N-myristoylation site"
 FT Modified-site 68..74
 FT /note= "N-myristoylation site"
 FT Region 77..80
 FT /note= "cell attachment sequence"
 FT Modified-site 216..222
 FT /note= "N-myristoylation site"

PN WO200037640-A2.

XX 29-JUN-2000.

XX 16-DEC-1999; 99MO-US30095.

XX 22-DEC-1998; 98US-0113296.

XX 08-MAR-1999; 99MO-US05028.

XX 02-JUN-1999; 99MO-US12252.

XX 01-SEP-1999; 99MO-US20111.

XX 15-SEP-1999; 99MO-US21090.

XX 30-NOV-1999; 99MO-US28313.

XX 01-DEC-1999; 99MO-US28409.

XX 02-DEC-1999; 99MO-US28301.

XX (GETH) GENENTECH INC.

XX Botstein D, Goddard A, Gurney AL, Hillan K, Lawrence DA, Roy MA;

XX Wood WI;

XX MPI: 2000-452188/39.

XX N-PSDB; AAA46907.

XX New anti-polypeptide antibody useful in the treatment and diagnosis of

XX neoplastic cell growth and proliferation

XX Claim 61: Fig 10; 220pp; English.

XX Sequence 243 AA:

Query Match 100.0%; Score 1325; DB 21; Length 243;

Best Local Similarity 100.0%; Pred. No. 5.3e-103; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTPGHNSOGLPGHGRDGRDGRGARG 60
 DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPCTPGHNSOGLPGHGRDGRDGRGARG 60
 QY 61 ARGEKGGGRPLPGRCGDPGRGEGPAGPTGPAECVPRPSAFSARSSSRVPPSPD 120
 DB 61 ARGEKGGGRPLPGRCGDPGRGEGPAGPTGPAECVPRPSAFSARSSSRVPPSPD 120
 QY 121 APLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATYRASLOFDLVKNGESIASFQ 180
 DB 121 APLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATYRASLOFDLVKNGESIASFQ 180
 QY 181 FFGWPKPASISGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLYSDMHSSP 240
 DB 181 FFGWPKPASISGAMVRLPEPDQVWVGVDYIGIYASIKTDTSTFSGFLYSDMHSSP 240

PA (GETH) GENENTECH INC.
XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
PI Gerritsen ME, Goddard A, Godowski P, Gurney AL, Sherwood S,
PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
XX MPI: 2001-408281/43.
DR N-PSDB: AAS21424.
XX
XX Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
XX lung, breast, prostate, cervical
XX
PS Claim 12: Flg 362; 813pp; English.
XX
XX AAU1217-AAU1246 represent novel human secretory and transmembrane
CC PRO polypeptides. The PRO polypeptides are useful to detect other
CC PRO polypeptides, to link bioactive molecules to cells expressing
CC PRO polypeptides, to modulate biological activities of cells expressing
CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
CC polypeptide expression in a cell sample to that in a control sample.
CC Some of the 275 sequences are also useful to stimulate the release of
CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
CC proliferation or differentiation of chondrocytes, the proliferation or
CC gene expression in pericyte cells, the release of proteoglycans from
CC cartilage, the proliferation of inner ear utricular supporting cells or
CC of T-lymphocytes, the release of a cytokine from peripheral blood
CC monocytes (PBMCs), or the proliferation of endothelial cells. Some of
CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
CC to factor VIIa. The PRO polypeptides can be used in assays to identify
CC molecules involved in binding interactions. The polynucleotides encoding
CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
CC transgenic or knock out animals and can be used in gene therapy.
XX
SQ Sequence 243 AA:
Query Match 100.0%; Score 1325; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 5.3e-103;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPLVLLLLGLAASPPDDNKIPSLCPGHPGLPCTPGHHSQGLPGDRGDRGARG 60
DB 1 MRPLVLLLLGLAASPPDDNKIPSLCPGHPGLPCTPGHHSQGLPGDRGDRGARG 60
QY 61 AGEKGEGRGRLPGPRGDPGRGEGAPGPTGPAEGCSVPPRSASFSAKRSRVPSPD 120
DB 61 AGEKGEGRGRLPGPRGDPGRGEGAPGPTGPAEGCSVPPRSASFSAKRSRVPSPD 120
QY 121 APLPDRVLVNOGHYDAVTGKFTCOVPGVYFAVAHATYRASLQFDLVKNESIASFFQ 180
DB 121 APLPDRVLVNOGHYDAVTGKFTCOVPGVYFAVAHATYRASLQFDLVKNESIASFFQ 180
QY 181 FFGGMPKPSLSGAMVRLPEPDQVWVGVDYIGIVASITDSTFSGFLVYSDMHSSP 240
DB 181 FFGGMPKPSLSGAMVRLPEPDQVWVGVDYIGIVASITDSTFSGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243
RESULT 9
AAB65815
ID AAB65815 standard; Protein: 243 AA.
XX
AC AAB65815;
XX
DT 28-MAR-2001 (first entry)
XX
DE Human TANGO 253 SEQ ID NO: 3.

XX Human: mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KW INTERCEPT 258; coronary disorder; olfactory disorder;
KW neurological disorder; pulmonary disorder; immunological disorder;
KW developmental disorder; kidney disorder.
XX
OS Homo sapiens.
PN WO200078808-A1.
XX
XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000MO-US16883.
XX
XX 18-JUN-1999; 99US-0336536.
XX
XX (MILL-) MILLENNIUM PHARM INC.
PI Leiby KR, McKay C, Bossone S;
DR MPI: 2001-050109/06.
XX
XX New nucleic acids for treating diseases and disorders, e.g.
PT atherosclerosis, infection, autoimmune diseases, obesity, ear
PT disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT sclerosis and asthma
XX
PS Claim 9; Page 211-212; 332pp; English.
XX
XX The present invention provides the protein and coding sequences of the
CC human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC coronary, pulmonary, olfactory, immunological, neurological,
CC developmental and kidney disorders.
XX
SQ Sequence 243 AA:
Query Match 100.0%; Score 1325; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 5.3e-103;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPLVLLLLGLAASPPDDNKIPSLCPGHPGLPCTPGHHSQGLPGDRGDRGARG 60
DB 1 MRPLVLLLLGLAASPPDDNKIPSLCPGHPGLPCTPGHHSQGLPGDRGDRGARG 60
QY 61 AGEKGEGRGRLPGPRGDPGRGEGAPGPTGPAEGCSVPPRSASFSAKRSRVPSPD 120
DB 61 AGEKGEGRGRLPGPRGDPGRGEGAPGPTGPAEGCSVPPRSASFSAKRSRVPSPD 120
QY 121 APLPDRVLVNOGHYDAVTGKFTCOVPGVYFAVAHATYRASLQFDLVKNESIASFFQ 180
DB 121 APLPDRVLVNOGHYDAVTGKFTCOVPGVYFAVAHATYRASLQFDLVKNESIASFFQ 180
QY 181 FFGGMPKPSLSGAMVRLPEPDQVWVGVDYIGIVASITDSTFSGFLVYSDMHSSP 240
DB 181 FFGGMPKPSLSGAMVRLPEPDQVWVGVDYIGIVASITDSTFSGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243
RESULT 10
AAB49593
ID AAB49593 standard; Protein: 243 AA.
XX
AC AAB49593;
XX
DT 13-MAR-2001 (first entry)
XX
XX Human adipocyte complement related protein homolog zs1g39.
DE Human; zacrps; gene therapy; complement inhibition; C1q domain;

KW adipocyte complement related protein homolog;
 KW inflammation; hormone secretion; inositol phosphate; arachidonate;
 KW phospholipase C activation; gastric emptying; neutrophil activation;
 KW superoxide anion production; antimicrobial; acute vascular injury;
 KW wound healing; zslg39.
 XX
 OS Homo sapiens.
 PN MO200073444-A1.
 PD 07-DEC-2000.
 PF 18-MAY-2000; 2000MO-US13608.
 PR 27-MAY-1999; 99US-0321372.
 PA (ZYMO) ZYMOGENETICS INC.
 PI Pliddington CS, Sheppard PO;
 DR WPI: 2001-061531/07.
 XX
 PT Novel adipocyte complement related protein homolog, ZACRP5, useful for
 PT diagnosing and treating inflammation, vascular injury microbial
 PT infections, and in wound healing
 PS Disclosure: Fig 1; 121pp; English.
 CC The present invention relates to human adipocyte complement related
 CC protein homolog, zacrp5 protein and coding sequence (see AAB49599 and
 CC AAC90045). ZACRP5 has a carboxyl-terminal C1q domain. The zacrp5 gene is
 CC located on human chromosome 16. zacrp5 gene and protein are useful for
 CC diagnosing and treating inflammations, for determining arterial
 CC remodelling, for modulating calcium ion concentration, hormone
 CC secretion, DNA synthesis or cell growth, inositol phosphate turnover,
 CC arachidonate release, phospholipase C activation, gastric emptying, human
 CC neutrophil activation or ADCC capability and superoxide anion production.
 CC zacrp5 gene and protein are also useful as antimicrobial applications,
 CC preferably against bacteria and virus, for complement inhibition, for
 CC treating acute vascular injury, and for wound healing. The present
 CC sequence is human adipocyte complement related protein homolog, zslg39
 CC protein. This protein was used in a sequence homology alignment with
 CC zacrp5 protein.
 XX
 SQ Sequence 243 AA:
 Query Match 100.0%; Score 1325; DB 22; Length 243;
 Best Local Similarity 100.0%; Pred. No. 5.3e-103;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 YY 1 MRRLVLLLLGLAAGSPRLDNNKIPSLCPGHPGLPTGPHHSGGLRGDRGRCAGP 60
 DB 1 MRRLVLLLLGLAAGSPRLDNNKIPSLCPGHPGLPTGPHHSGGLRGDRGRCAGP 60
 YY 61 APGKGGGGRPGGLPGRPGDPRGEGAPRPTGAGCSVPPRSASFSAKSESSEVPPSPD 120
 DB 61 APGKGGGGRPGGLPGRPGDPRGEGAPRPTGAGCSVPPRSASFSAKSESSEVPPSPD 120
 YY 121 APLPFDRVLVNEQGHDAVTGKFTCOVPGYVYFAVNAHYRASLQDFLVKNGESIASFFQ 180
 DB 121 APLPFDRVLVNEQGHDAVTGKFTCOVPGYVYFAVNAHYRASLQDFLVKNGESIASFFQ 180
 YY 181 FFGGMPKRPASLSGAMVRLEREDQVWVQVGVGYIGIYASIKTDSFGFLVYSDMHSSP 240
 DB 181 FFGGMPKRPASLSGAMVRLEREDQVWVQVGVGYIGIYASIKTDSFGFLVYSDMHSSP 240
 YY 241 VFA 243
 DB 241 VFA 243
 RESULT 11
 AAB49599

ID AAB49599 standard; Protein: 243 AA.
 XX
 AC AAB49599;
 XX
 DT 13-MAR-2001 (first entry)
 DE
 XX Human adipocyte complement related protein homolog zslg39.
 KW Human; zacrp6; gene therapy; complement inhibition; C1q domain;
 KW adipocyte complement related protein homolog;
 KW inflammation; hormone secretion; inositol phosphate; arachidonate;
 KW phospholipase C activation; gastric emptying; neutrophil activation;
 KW superoxide anion production; antimicrobial; acute vascular injury;
 KW wound healing; zslg39.
 XX
 OS Homo sapiens.
 PN MO200073446-A2.
 PD 07-DEC-2000.
 PF 22-MAY-2000; 2000MO-US14024.
 PR 27-MAY-1999; 99US-0321262.
 PA (ZYMO) ZYMOGENETICS INC.
 PI Pliddington CS, Sheppard PO;
 DR WPI: 2001-061532/07.
 XX
 PT Novel adipocyte complement related protein homolog, ZACRP6, useful as
 PT modulators of neurotransmission and for treating disseminated
 PT intravascular coagulation, arteriosclerosis and acute vascular injury
 PS Disclosure: Fig 1; 119pp; English.
 CC The present invention relates to human adipocyte complement related
 CC protein homolog, ZACRP6 protein and coding sequence (see AAB49596 and
 CC AAC90051). ZACRP6 has a carboxyl-terminal C1q domain. The zacrp6 gene is
 CC located on human chromosome 21q. zacrp6 gene and protein are useful for
 CC diagnosing and treating inflammations, for determining arterial
 CC remodelling, for modulating calcium ion concentration, hormone
 CC secretion, DNA synthesis or cell growth, inositol phosphate turnover,
 CC arachidonate release, phospholipase C activation, gastric emptying, human
 CC neutrophil activation or ADCC capability and superoxide anion production.
 CC zacrp6 gene and protein are also useful as antimicrobial applications,
 CC preferably against bacteria and virus, for complement inhibition, for
 CC treating acute vascular injury, disseminated intravascular coagulation,
 CC arteriosclerosis and for wound healing. The present sequence is human
 CC adipocyte complement related protein homolog zslg39. This protein was
 CC used in a sequence homology comparison with ZACRP6 protein.
 XX
 SQ Sequence 243 AA:
 Query Match 100.0%; Score 1325; DB 22; Length 243;
 Best Local Similarity 100.0%; Pred. No. 5.3e-103;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 YY 1 MRRLVLLLLGLAAGSPRLDNNKIPSLCPGHPGLPTGPHHSGGLRGDRGRCAGP 60
 DB 1 MRRLVLLLLGLAAGSPRLDNNKIPSLCPGHPGLPTGPHHSGGLRGDRGRCAGP 60
 YY 61 APGKGGGGRPGGLPGRPGDPRGEGAPRPTGAGCSVPPRSASFSAKSESSEVPPSPD 120
 DB 61 APGKGGGGRPGGLPGRPGDPRGEGAPRPTGAGCSVPPRSASFSAKSESSEVPPSPD 120
 YY 121 APLPFDRVLVNEQGHDAVTGKFTCOVPGYVYFAVNAHYRASLQDFLVKNGESIASFFQ 180
 DB 121 APLPFDRVLVNEQGHDAVTGKFTCOVPGYVYFAVNAHYRASLQDFLVKNGESIASFFQ 180
 YY 181 FFGGMPKRPASLSGAMVRLEREDQVWVQVGVGYIGIYASIKTDSFGFLVYSDMHSSP 240

Search completed: June 18, 2003, 14:51:05
Job time : 87 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 14:41:33 ; Search time 81 Seconds
(without alignments)
618.141 Million cell updates/sec

Title: US-09-943-851A-42
Perfect score: 1325
Sequence: 1 MRPLLVLLLLGLAGSPPLD.....DSTFGFLVSDWHSPPYFA 243

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues
Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_protist:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriophage:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1254	94.6	243	11	08R002
2	435.5	32.1	240	6	Q95M04
3	422	31.8	243	6	Q95J07
4	415.5	31.4	675	6	Q9N178
5	402	30.3	295	11	Q921K4
6	395	29.8	294	11	Q9D8U4
7	373.5	28.2	194	6	Q95J35
8	370.5	28.0	744	11	Q9D2V4
9	370.5	28.0	744	11	Q921S8
10	368	27.8	744	4	Q96D07
11	363.5	27.4	705	4	Q9TEJ5
12	314	23.7	245	11	Q9DCM6
13	270.5	20.4	246	11	Q9ES30
14	266.5	20.1	196	11	Q920N0
15	252.5	19.1	246	13	Q91907
16	247.5	18.7	256	13	Q91909

17	241.5	18.2	347	4	Q961H6
18	238.5	18.0	583	4	Q96G58
19	238.5	18.0	992	4	Q9UG76
20	238.5	18.0	1016	4	Q9Y6C2
21	237	17.9	890	5	Q77087
22	235	17.7	1017	11	Q99K41
23	234.5	17.7	173	6	Q62789
24	225	17.0	325	5	Q17036
25	223	16.8	281	11	Q9QXP7
26	222.5	16.8	322	5	Q01945
27	222	16.8	251	13	Q91908
28	222	16.8	1378	5	Q97405
29	221.5	16.7	120	6	Q77782
30	221.5	16.7	248	6	Q9T706
31	220	16.6	341	5	Q20142
32	219.5	16.6	205	11	Q9D0W2
33	219.5	16.6	326	11	Q8R066
34	219.5	16.6	381	5	Q94399
35	219.5	16.6	589	11	Q99L16
36	219.5	16.6	1453	11	Q63079
37	218	16.5	182	11	Q8R1P2
38	217	16.4	319	5	Q17038
39	217	16.4	1160	4	Q14046
40	217	16.4	1344	13	Q93419
41	217	16.4	1418	6	Q28396
42	217	16.4	1442	11	Q62031
43	217	16.4	1442	11	Q62033
44	217	16.4	1450	13	Q9Y1B4
45	217	16.4	1459	11	Q62032

ALIGNMENTS

RESULT 1

Q8R002 PRELIMINARY: PRT: 243 AA.
ID Q8R002
AC Q8R002
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to DKE2P586B0621 protein (Hypothetical 25.4 kDa protein).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC023068; AAH23068.1;
DR EMBL; BC025174; AAH25174.1;
KW Hypothetical protein.
SQ SEQUENCE 243 AA; 25420 MW; 498129CD051DB97B CRC64;

Query Match 94.6%; Score 1254; DB 11; Length 243;
Best Local Similarity 93.4%; Pred. No. 5.1e-100;
Matches 227; Conservative 7; Mismatches 9; Indels 0; Gaps 0;
QY 1 MRPLLVLLLLGLAGSPPLDNRKIPSLCPGHPGLPTGHHGSGCLPGRGDRDGDAGP 60
DB 1 MRPLLVLLLLGLVSGSPPLDNRKIPSLCPGHPGLPTGHHGSGCLPGRGDRDGDAGP 60
QY 61 APEKGGGRRGRLPGPRGDRGPRGACRPACPTGACGCSVPNSAFSAKSESRRPPSD 120
DB 61 APEKGGGRRGRLPGPRGDRGPRGACRPACPTGACGCSVPNSAFSAKSESRRPPSD 120
QY 121 APLPFDRLVNEQGHYAVGKFTCOVGYVYFAVNAVTVRASLPGLVNGESIASFFQ 160
DB 121 APLPFDRLVNEQGHYAVGKFTCOVGYVYFAVNAVTVRASLPGLVNGESIASFFQ 160

Db 121 TPLEDRVLTNEGHPPTGKFTCOYVGVYFAVHATVYASIQFDLVKNGSIASFQ 180
 QY 181 FFGWMPKPRASISGAMVRLPEDEQVWVGVGVYIGIYASIKTSTSGFLVSDWHSPP 240
 Db 181 YFGWMPKPRASISGAMVRLPEDEQVWVGVGVYIGIYASIKTSTSGFLVSDWHSPP 240
 QY 241 VFA 243
 Db 241 VFA 243
 RESULT 2
 095M04 PRELIMINARY: PRT: 240 AA.
 AC 095M04: PRELIMINARY: PRT: 240 AA.
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Adipose tissue-specific protein adipo Q.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovinae; Bos.
 NCBI_Taxid=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21359933; PubMed-11382781;
 RA Sato C., Yasukawa Z., Honda N., Matsuda T., Kitaajima K.;
 RT Identification and Adipocyte Differentiation-dependent Expression of
 the Unique Distal Acid Residue in an Adipose Tissue-specific
 Glycoprotein, Adipo Q.
 J. Biol. Chem. 276:28849-28856(2001).
 RL EMBL: AF269230; AAK58902.1; -.
 DR Interpro: IPR001073; Clq.
 DR Interpro: IPR000087; Collagen.
 DR Pfam: PF00386; Clq; 1.
 DR Pfam: PF01391; Collagen; 1.
 DR Prodom: PD000007; Collagen; 1.
 DR PROSITE: PS01113; Clq; UNKNOWN_1.
 SO SEQUENCE 240 AA: 26091 MW: C6253BA80389A668 CRC64;
 Query Match 32.1%; Score 425.5; DB 6; Length 240;
 Best Local Similarity 41.5%; Pred. No. 7,1e-29;
 Matches 100; Conservative 35; Mismatches 81; Indels 25; Gaps 10;
 QY 6 VLLILGLAAGSPPLDNDKIP-SLCPG-HPGRLPFGHHSOGILPGRDGRDGAAGAP 62
 Db 7 LLLLLLPSSHGDNMDPRLPKGACAGMAGIRGHPGHNCT---PGRDGRD-----GTP 57
 QY 63 GEKGEGRGLPGRPD-----PGRGEAGPAGCTPAGCSVPPRSASFSAKSES 113
 Db 58 GEKGEKGDAGLLGPKETGCVGTGAEGRPGTGRKGEPEAAVYRSASFV-GLFT 116
 QY 114 RVPPSDAPLPEDRVLTNEGHPPTGKFTCOVGVYFAVHATVYASIQFDLVKNG 173
 Db 117 RVTVVP-NVPIRFTKIFYNQNHHDGSGTKYCNIPGLYFHYITVMKVKSLRKKDK 175
 QY 174 STASFOFGGMPKPRASISGAMVRLPEDEQVWVGVGVYIGIYASIKTSTSGFLV 232
 Db 176 AVLEFYDQYOE-KNVDAQSGSVLLHLEVGDQVWLQYEGENHNGVYADANNDSFTGFL 234
 QY 233 Y 233
 Db 233 Y 233
 RESULT 3
 095J07 PRELIMINARY: PRT: 243 AA.
 AC 095J07: PRELIMINARY: PRT: 243 AA.
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)

DE Adiponectin.
 GN APMI.
 OS Macaca mulatta (Rhesus macaque).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
 OC Cercopithecoidea; Macaca.
 NCBI_Taxid=9544;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-ADIPOSE TISSUE:
 RX MEDLINE-21232234; PubMed-11334417;
 RA Hotta K., Funahashi T., Bodkin N.L., Ohtmeier H.K., Arlia Y.,
 Hansen B.C., Matsuzawa Y.;
 RT "Circulating concentrations of the adipocyte protein adiponectin are
 decreased in parallel with reduced insulin sensitivity during the
 progression to type 2 diabetes in rhesus monkeys."
 Diabetes 50:1126-1133(2001).
 RL EMBL: AF044407; AAK92202.1; -.
 DR Interpro: IPR001073; Clq.
 DR Interpro: IPR000087; Collagen.
 DR Pfam: PF00386; Clq; 1.
 DR Pfam: PF01391; Collagen; 1.
 DR Prodom: PD000007; Collagen; 1.
 DR PROSITE: PS01113; Clq; UNKNOWN_1.
 SO SEQUENCE 243 AA: 26264 MW: 49A45DAF2B4613PD CRC64;
 Query Match 31.8%; Score 422; DB 6; Length 243;
 Best Local Similarity 42.3%; Pred. No. 1,4e-28;
 Matches 104; Conservative 28; Mismatches 88; Indels 26; Gaps 8;
 QY 6 VLLILGLAAGSPPLDNDKIPSLCPGHPG-----LPGRGHHHSOGILPGRDGRDGAAG 60
 Db 6 VLLILLPSSHGODPTTQGPVLLPLPKAGCTGMAGIRGHPGHNQVGRDGRD-----G 59
 QY 61 APGEKGEGRGLPGRPD-----PGRGEAGPAGCTPAGCSVPPRSASFSAKRS 111
 Db 60 TPGEKGEKGDGPGGLGPRGDTGEGVTGAEGRPGTGRKGEPEAAVYRSASFV-GL 118
 QY 112 ESRVPPSDAPLPEDRVLTNEGHPPTGKFTCOVGVYFAVHATVYASIQFDLVKN 171
 Db 119 ETVTVVP-NMPIRFTKIFYNQNHHDGSGTKYCNIPGLYFHYITVMKVKSLRKKDK 177
 QY 172 GES-IASFQFGGMPKPRASISGAMVRLPEDEQVWVGVGVYIGIYASIKTSTSG 229
 Db 178 DKAMLEFYDQYOEENNDQA---SSVLLHLEVGDQVWLQYEGEGRGLVADANNDSFTG 235
 QY 230 FLYVSD 235
 Db 236 FLYVHD 241
 RESULT 4
 09N178 PRELIMINARY: PRT: 675 AA.
 AC 09N178: PRELIMINARY: PRT: 675 AA.
 DT 01-OCT-2000 (TREMBlrel. 15, Created)
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Type X collagen.
 GN COL10A1.
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.
 NCBI_Taxid=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-21015405; PubMed-11130976;
 RA Nielsen V.H., Bendixen C., Arnbjerg J., Sorensen C.M., Jensen H.E.,
 Shukri N.M., Thomsen B.;
 RT "Abnormal growth plate function in pigs carrying a dominant mutation
 in type X collagen."
 Mamm. Genome 11:1087-1092(2000).
 RL EMBL: AF222861; AAF37271.1; -.

QY 3 PLLVLLLLAGSPPDLNKKIPSLCPHGLPRTGPHHSGCLGRDGRDGRDGPAP 62
 DB 41 POLVCLSP------POGPPGPPGAGSSGVYGRMGFPKGGODGODGODGRDS 87
 QY 63 GEGEGEGRP-GLPGRGDPGRGEGAPG-TPAGDECSVPRPSAKRESRKP 116
 DB 88 GEGEGEGRGNGKNGKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQKQK 147
 QY 117 PPSDA-----PLPDRVLNVEGCHDAYTGKFTCOVPGVYFA 154
 DB 148 GPCSCSSSHAKSAFSAVATKSYPRERLPKFKEDKIILNBSGHVNASCKFVCSPPCIYFT 207
 QY 155 VYATYRASLOPDLVKNNGESISFQFGGKPKRPSLSGAVNLEPDDQVWVGVGDY 214
 DB 208 VQITLANKHLAGLVHNGQYRTRTDANTGNHDA--SGSTILALKEDEWLOIFYSQ 265
 QY 215 IGI-YASIKTSTFSGFLVYSD 235
 DB 266 NGLFYDPYWTDSLTFGFLIYAD 287

RESULT 7
 ID 095J95 PRELIMINARY; PRT: 194 AA.
 AC 095J95;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Adiponectin (Fragment).
 GN APM1.
 OS Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 OX NCBI_TaxID=9615;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=ADIPOSE TISSUE;
 RA Kabir M., Anantharayan S., Ionut V., Kim S.P., Van Citters G.W.,
 RA Dea M.K., Bergman R.N.;
 RL "Regulation of Adiponectin gene expression in the fat-fed dog";
 RL Submitted (Sep-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF417206; AAL09702.1;
 DR InterPro: IPR001073; Clq.
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF00386; Clq; 1.
 DR Pfam: PF01391; Collagen; 1.
 DR ProDom: PD000007; Collagen; 1.
 DR ProSite: PS01113; Clq; UNKNOWN_1.
 FT NON_TER 1
 FT 194
 SQ SEQUENCE 194 AA; 20890 MW; 3AA3D947D187AF9A CRC64;

Query Match 28.2%; Score 373.5; DB 6; Length 194;
 Best Local Similarity 42.3%; Pred. No. 1.6e-24;
 Matches 90; Conservative 22; Mismatches 66; Indels 35; Gaps 9;

QY 8 LLLGLAGSPRLDNKIPSLCPG-HRCLPRTGPHHSGCLGRDGRDGRDGPAP 66
 DB 4 VLLPRLPGA-----CRGMMAGIRGHPGNGT--PGRGRD-----GTPEKG 43
 QY 67 EGRNRPGLRGRD-----PGRGEGAPRAGTGPAGCSVPRPSAKRESRKP 117
 DB 44 EKGRPGGLVGRKGTGVTGVEPRGRFRTGKRGKRGESATVYHNSAFV--GLESRTIV 102
 QY 118 PPSDAPLPDRVLNVEGCHDAYTGKFTCOVPGVYFAVYATYRASLOPDLVKNNGESIAS 177
 DB 103 P-WNPRIFTFIYFNLDNHVDTGCKFKCNIRGLYFESYHITVYLKDYKVSLLYKDK--AM 159
 QY 178 FQFQFGMFKRA--SLSGAVNLEPDDQVWV 209
 DB 160 LFTYDQYQEKVNDQASGVLLHLLEVGDQWLDY 192

RESULT 8
 ID 09D2V4 PRELIMINARY; PRT: 744 AA.
 AC 09D2V4;
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Procollagen, type VIII, alpha 1.
 GN COL8A1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamaneke I.,
 RA Saito T., Okazaki Y., Gotohori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fletschmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
 RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszewski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.;
 RT "Functional annotation of a full-length mouse cDNA collection."
 RL Nature 409:685-690(2001).
 DR EMBL: AK018742; BAB31383.1;
 DR MCD: MGI:88463; Col8a1.
 DR InterPro: IPR001073; Clq.
 DR InterPro: IPR000087; Collagen.
 DR Pfam: PF00386; Clq; 1.
 DR Pfam: PF01391; Collagen; 7.
 DR PRINTS: PR00007; COMPLEMENTC1Q.
 DR SMART: SM00110; Clq; 1.
 DR ProSite: PS01113; Clq; 1.
 SQ SEQUENCE 744 AA; 73581 MW; C659BDCBCB06EB9C CRC64;

Query Match 28.0%; Score 370.5; DB 11; Length 744;
 Best Local Similarity 34.6%; Pred. No. 1.4e-23;
 Matches 93; Conservative 32; Mismatches 83; Indels 61; Gaps 7;

QY 24 IPRSLC--PGRNRPGLRGRHNGSCLGRDGRDGRDGRDGPAPGKGGKGRGLP----- 74
 DB 476 VPGLLPKRGEGIPRGDGLGRPGIRGIVGSPRIPRGPGRKGRGLRGRGFGVGR 535
 QY 75 -----GPRKDPGRGEGAPRAGTGP-----AGCSV 100
 DB 536 PGVAGLHPRPKRGALGRGQGRLPGRPRGRGRRAVMPTRPSGEGVLPDMGLGIDVY 595
 QY 101 PPSASASAKR-----SSRVP-PPSDAPLPDRVLNVEGCHDAYTGKFTC 145
 DB 596 KPRHAYAGKKGKNGKGRPAVEMAFATLALVFPFVCAVFKDLKLYGRNVPQICIFTC 655
 QY 146 QVGVYTFVFNHTVYRASLOPDLVKNNGESIA-SFQFGGKPKRPSLSGAVNLEPDDQ 204
 DB 656 EYRGVYTFVFNHCKGKGNVWALFKNNRPMYTYDYKGFIDQA--SGSAVLLLRPDQ 713
 QY 205 VVWQVGVGDYIGIYASIKTSTFSGFLVY 233
 DB 714 VFLQMPSEQAGLAVAGQVYVHSSFSGLVY 742

DB 643 NWPATYTYDEXKGYLDA--SGGAVLQLRPNQVWVQMPSDQANGLYSTREYTHSSRSGF 700
 OY 231 LV 232
 Db 701 LL 702

RESULT 12

09DCM6 PRELIMINARY: PRT: 245 AA.
 AC 09DCM6:
 DT 01-JUN-2001 (TREMBlrel. 17, Created)
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
 DE Complement component 1, q subcomponent, alpha polypeptide.
 CN C10A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=KIDNEY;
 RX MEDLINE=21085660; PubMed=11217851;
 RA Kawal J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 RA Alzawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka S.,
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiya H.,
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
 RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
 RA Brownstein M.J., Bult C., Fletcher M., Fujita M., Gariboldi M.,
 RA Gustincich S., Hall D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombauts P.,
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 RA Sasaki H., Sato K., Schoenwald M., Rodriguez I., Sakamoto N.,
 RA Suzuki H., Togo-Oka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 RA Wyszynski A., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
 RA Hayashizaki Y.,
 RT *Functional annotation of a full-length mouse cDNA collection.*;
 RN Nature 409:685-690(2001).
 RN [2]

RP SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK002655; BAB22262.1;
 DR EMBL: BC002086; AA02086.1;
 DR MGD: MGI:88223; C1qa.
 DR Interpro: IPR001073; C1q.
 DR Interpro: IPR000087; Collagen.
 DR Pfam: PF00386; C1q; 1.
 DR Pfam: PF01391; Collagen; 1.
 DR PRINTS: PRO0007; COMPLEMENTC1Q.
 DR SMART: SM00110; C1Q; 1.
 DR PROSITE: PS01113; C1Q; 1.
 SQ SEQUENCE 245 AA; 25974 MW; 41C2066D49592020 CRC64;

Query Match 23.74; Score 314; DB 11; Length 245;
 Best Local Similarity 35.38; Pred. No. 2.7e-19;
 Matches 79; Conservative 24; Mismatches 85; Indels 36; Gaps 6;

OY 29 PGHGLPTPTGHHGSGQLPGRDGRDAPGAPGEGGGRGRLPGRDGRGEGR 88
 Db 36 PGHGLPTPTGHHGSGQLPGRDGRDAPGAPGEGGGRGRLPGRDGRGEGR 93
 OY 89 AGPTGAGGECV-----PRSAFSAKRSERVRPPSDAPLPEDRYLVNCGHYDAVTGK 142
 Db 94 SGGQGLKGVKGNKGNTRDPRPAFSAIRON---PMTLGNVVLFDKLTNGESPRYNHTR 150
 OY 143 FTGCVGVGVYFAVHATVYRASLQFDLVKNGESIASFFGFGWPKRA----- 189

DB 151 FICAVGGEYF-----NFQVSKMDLCLEIKSSSGGGRDLSSTNTNKKLQF 199
 OY 190 SLSSGAMRLPEPDVWVQVGVGYDTIYASIKRTDSTFGFLVY 233
 Db 200 VLAGSTVLQLRGDEWIEKDPANG-RYQCTEADSIFFGFLIF 242

RESULT 13

09ES30 PRELIMINARY: PRT: 246 AA.
 AC 09ES30:
 DT 01-MAR-2001 (TREMBlrel. 16, Created)
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
 DE Collagenous repeat-containing sequence of 26kDa protein.
 CN COR3 OR COR326.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21264842; PubMed=11071891;
 RA Maeda T., Abe M., Kurisu K., Jikko A., Furukawa S.;
 RT "Molecular cloning and characterization of a novel gene, COR326,
 RT encoding a putative secretory protein and its possible involvement in
 RT skeletal development";
 RL J. Biol. Chem. 276:3628-3634(2001).
 DR EMBL: AF246265; AAC33704.1;
 DR MGD: MGI:1932136; COR3.
 DR Interpro: IPR001073; C1q.
 DR Interpro: IPR000087; Collagen.
 DR Pfam: PF00386; C1q; 1.
 DR Pfam: PF01391; Collagen; 1.
 DR PRINTS: PRO0007; COMPLEMENTC1Q.
 DR SMART: SM00110; C1Q; 1.
 DR PROSITE: PS01113; C1Q; 1.
 SQ SEQUENCE 246 AA; 26828 MW; 42A481B3E9F48F7E CRC64;

Query Match 20.48; Score 270.5; DB 11; Length 246;
 Best Local Similarity 31.88; Pred. No. 1.5e-15;
 Matches 78; Conservative 33; Mismatches 107; Indels 27; Gaps 6;

OY 4 LVLVLLGLA-----AGSPPLDNRKTPSLCPGHPGLPTGPHGSGQLPGRDGR 52
 Db 11 LVLVLLPPLCLQDDEYMSPPQAGLPPDSCSKCHGDYGRFYOGPPGPPGIPGNHGN 70
 OY 53 DGRDAPGAPGEGGGRGRLPGRDGRDAPGAPGEGGGRGRLPGRDGRGEGR 112
 Db 71 NNNNATGHEGAKGE-----KGDKGDLGPRGEROHGPKKEKTPVPPQLQIAFNASL 124
 OY 113 SRVPPSDAPLPEDVVLVNEGCHDAVTGKTCQVPGVYFAV---HATVYRASLQFDL 168
 Db 125 ATHESNNSGILFESSVEITNINFDVMTGREGAPVSGVYFFTFSMKHEDVEEYVY--L 182
 OY 169 VKNGESIASFFGFGWPKPASLSGAMVRLPEPDVWVQVGVGYDTIYASITDSTFS 228
 Db 183 MHNGTIVSMVSYELKKGSDTS-SNHAVLKLAKDEWILRMNGN---ALHGDHQRSTFA 238
 OY 229 GFLVY 233
 Db 239 GFLLE 243

[illegible]

RESULT	15
091907	
ID	091907
AC	091907; PRELIMINARY; PRT; 246 AA.
DT	01-OCT-2000 (TrEMBLrel. 15, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT	01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE	Mannose binding-like lectin precursor (Fragment).
GN	MBL.
OS	Carassius auratus (Goldfish).
OC	Euraryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC	Cyprinidae; Carassius.
OX	NCBI_TaxID=7957;
RN	[1]
RP	SEQUENCE FROM N.A.
RC	TISSUE=LIVER;
RX	MEDLINE=20456722; PubMed=11003389;
RT	Vitved L., Holmskov U., Koch C., Telsner B., Hansen S., Skjoldt K.;
RT	"The homologue of mannose-binding lectin in the carp family Cyprinidae
RT	is expressed at high level in spleen, and the deduced primary
RL	structure predicts affinity for galactose.";
DR	EMBL: AF227739; AAF63470.1; -
DR	HSSP: P35247; 1B08.
DR	InterPro: IPR000087; Collagen
DR	InterPro: IPR001304; Lectin_C.
DR	Pfam: PF01391; Collagen; 2.
DR	Pfam: PR00059; Lectin_c; 1.
DR	SMART: SM00034; CLECT; 1.

Query Match	19.18;	Score 252.5;	DB 13;	Length 246
Best Local Similarity	31.98;	Pred. No. 5.2e-14;		
Matches 80;	Conservative 33;	Mismatches 73;	Indels 65	

```

QY      4 LTVLLILGLAASPPDLNDONKTIPLCLCPHGLPCTPGHHSQGLPGNDGHDGDAFAGAPG 63
Db      2 LILQFLALQLLDIAEHPQNLN-----CPAVGCGPGTGGHIN--GLPGDGDGDGDAIGPKG 53

QY      64 EKGEG-----GHPGLGPRGDPGPRGEGAGPAGTGPAGSECVPPRASFSAKRSSESRVPP 118
Db      54 EKGESGVSVYQGPCKAGPPGTAGEKGEGSPGQSPGSGSVLEISKSTIQLLKKTI-- 110

QY      119 SDAPLEPDRVLVNEGCHDAYTGKFTQVQVGYTFEAVHATVYRASLQDFLVNKGESIASF 178
Db      111 -----ATFER--VSSVCHERKV-----GQRYITTDGVG--NFD-----GKLKSC 146

QY      179 FQFFGGMKPASLSGAMV--RLEPEDVWVWY-----GVGD--YIGIYASIKTIDSTP-- 227
Db      147 MEF-----CGTMVSPRPSAENQALLKLVSSGLSKSKPKYIGV--TDKRTKEQFVD 194

QY      228 ---SGFLVYSDW 236
Db      195 TEGKQLTFETNM 205

```

```
Search completed: June 18, 2003, 14:52:31
Job time : 83 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 14:53:08 ; Search time 49 Seconds
(without alignments)
536.617 Million cell updates/sec

Title: US-09-943-851A-42

Sequence: 1 MRPLVLLGLLAGSPPLD.....DSRFSGFLVYSMDHSSPVEA 243

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 417779 seqs, 108206813 residues

Total number of hits satisfying chosen parameters: 417779

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1325	100.0	243	9	US-09-944-413-42
2	1325	100.0	243	9	US-09-944-403-42
3	1325	100.0	243	9	US-09-944-896-42
4	1325	100.0	243	9	US-09-944-944-42
5	1325	100.0	243	9	US-09-944-907-42
6	1325	100.0	243	9	US-09-944-929-42
7	1325	100.0	243	9	US-10-028-072-362
8	1325	100.0	243	9	US-10-121-049-362
9	1325	100.0	243	9	US-10-123-904-362
10	1325	100.0	243	9	US-10-140-470-362
11	1325	100.0	243	9	US-09-796-753-68
12	1325	100.0	243	9	US-10-175-746-362
13	1325	100.0	243	9	US-10-176-918-362
14	1325	100.0	243	9	US-10-137-865-362
15	1325	100.0	243	9	US-10-140-474-362
16	1325	100.0	243	9	US-10-142-431-362
17	1325	100.0	243	9	US-10-143-114-362
18	1325	100.0	243	9	US-10-140-002-362
19	1325	100.0	243	9	US-10-142-419-362

20	1325	100.0	243	9	US-10-123-262-362	Sequence 362, App
21	1325	100.0	243	9	US-10-142-423-362	Sequence 362, App
22	1325	100.0	243	9	US-10-142-423-362	Sequence 362, App
23	1325	100.0	243	9	US-10-141-755-362	Sequence 362, App
24	1325	100.0	243	9	US-10-141-755-362	Sequence 362, App
25	1325	100.0	243	9	US-10-143-032-362	Sequence 362, App
26	1325	100.0	243	9	US-10-123-108-362	Sequence 362, App
27	1325	100.0	243	9	US-10-123-236-362	Sequence 362, App
28	1325	100.0	243	9	US-10-123-261-362	Sequence 362, App
29	1325	100.0	243	9	US-10-140-921-362	Sequence 362, App
30	1325	100.0	243	9	US-10-140-921-362	Sequence 362, App
31	1325	100.0	243	9	US-10-121-045-362	Sequence 362, App
32	1325	100.0	243	9	US-10-123-292-362	Sequence 362, App
33	1325	100.0	243	9	US-10-123-903-362	Sequence 362, App
34	1325	100.0	243	9	US-10-124-819-362	Sequence 362, App
35	1325	100.0	243	9	US-10-124-822-362	Sequence 362, App
36	1325	100.0	243	9	US-10-160-498-362	Sequence 362, App
37	1325	100.0	243	9	US-09-944-884-42	Sequence 42, App
38	1325	100.0	243	9	US-10-121-041-362	Sequence 362, App
39	1325	100.0	243	9	US-10-121-041-362	Sequence 362, App
40	1325	100.0	243	9	US-10-121-041-362	Sequence 362, App
41	1325	100.0	243	9	US-10-121-047-362	Sequence 362, App
42	1325	100.0	243	9	US-10-123-215-362	Sequence 362, App
43	1325	100.0	243	9	US-10-123-908-362	Sequence 362, App
44	1325	100.0	243	9	US-10-123-908-362	Sequence 362, App
45	1325	100.0	243	9	US-10-123-909-362	Sequence 362, App

ALIGNMENTS

RESULT 1

US-09-944-413-42

Sequence 42, Application US/09944413

Patent No. US20020156004A1

GENERAL INFORMATION:

APPLICANT: Baker, Kevin

APPLICANT: Batstein, David

APPLICANT: Eaton, Dan

APPLICANT: Ferrara, Napoleone

APPLICANT: Filvaroff, Ellen

APPLICANT: Gerritsen, Mary

APPLICANT: Goddard, Audrey

APPLICANT: Grimaldi, Christopher

APPLICANT: Gurney, Austin

APPLICANT: Hillan, Kenneth

APPLICANT: Kijavlin, Ivar

APPLICANT: Napier, Mary

APPLICANT: Roy, Margaret

APPLICANT: Thomas, Daniel

APPLICANT: Wood, William

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

FILE REFERENCE: P2548P101

CURRENT APPLICATION NUMBER: US/09/944,413

CURRENT FILING DATE: 2001-09-26

PRIOR APPLICATION NUMBER: 09/866,028

PRIOR FILING DATE: 2001-05-25

PRIOR APPLICATION NUMBER: 60/067,411

PRIOR FILING DATE: December 3, 1997

PRIOR APPLICATION NUMBER: 60/069,334

PRIOR FILING DATE: December 11, 1997

PRIOR APPLICATION NUMBER: 60/069,335

PRIOR FILING DATE: December 11, 1997

PRIOR APPLICATION NUMBER: 60/069,278

PRIOR FILING DATE: December 11, 1997

PRIOR APPLICATION NUMBER: 60/069,425

PRIOR FILING DATE: December 12, 1997

PRIOR APPLICATION NUMBER: 60/069,696

PRIOR FILING DATE: December 16, 1997

PRIOR APPLICATION NUMBER: 60/069,694

PRIOR FILING DATE: December 16, 1997

PRIOR APPLICATION NUMBER: 60/069,702
 PRIOR FILING DATE: December 16, 1997
 PRIOR APPLICATION NUMBER: 60/069,870
 PRIOR FILING DATE: December 17, 1997
 PRIOR APPLICATION NUMBER: 60/069,873
 PRIOR FILING DATE: December 17, 1997
 PRIOR APPLICATION NUMBER: 60/068,017
 PRIOR FILING DATE: December 18, 1997
 PRIOR APPLICATION NUMBER: 60/070,440
 PRIOR FILING DATE: January 5, 1998
 PRIOR APPLICATION NUMBER: 60/074,086
 PRIOR FILING DATE: February 9, 1998
 PRIOR APPLICATION NUMBER: 60/074,092
 PRIOR FILING DATE: February 9, 1998
 PRIOR APPLICATION NUMBER: 60/075,945
 PRIOR FILING DATE: February 25, 1998
 PRIOR APPLICATION NUMBER: 60/112,850
 PRIOR FILING DATE: December 16, 1998
 PRIOR APPLICATION NUMBER: 60/113,296
 PRIOR FILING DATE: December 22, 1998
 PRIOR APPLICATION NUMBER: 60/146,222
 PRIOR FILING DATE: July 28, 1999
 PRIOR APPLICATION NUMBER: PCT/US98/19330
 PRIOR FILING DATE: September 16, 1998
 PRIOR APPLICATION NUMBER: PCT/US98/25108
 PRIOR FILING DATE: December 1, 1998
 PRIOR APPLICATION NUMBER: 09/216,021
 PRIOR FILING DATE: December 16, 1998
 PRIOR APPLICATION NUMBER: 09/218,517
 PRIOR FILING DATE: December 22, 1998
 PRIOR APPLICATION NUMBER: 09/254,311
 PRIOR FILING DATE: March 3, 1999
 PRIOR APPLICATION NUMBER: PCT/US99/12252
 PRIOR FILING DATE: June 22, 1999
 PRIOR APPLICATION NUMBER: PCT/US99/21090
 PRIOR FILING DATE: September 15, 1999
 PRIOR APPLICATION NUMBER: PCT/US99/28409
 PRIOR FILING DATE: No. US20020156004A1e1ember 30, 1999
 PRIOR APPLICATION NUMBER: PCT/US99/28313
 PRIOR FILING DATE: No. US20020156004A1e1ember 30, 1999
 PRIOR APPLICATION NUMBER: PCT/US99/28301
 PRIOR FILING DATE: December 1, 1999
 PRIOR APPLICATION NUMBER: PCT/US99/30095
 PRIOR FILING DATE: December 16, 1999
 PRIOR APPLICATION NUMBER: PCT/US00/03565
 PRIOR FILING DATE: February 11, 2000
 PRIOR APPLICATION NUMBER: PCT/US00/04414
 PRIOR FILING DATE: February 22, 2000
 PRIOR APPLICATION NUMBER: PCT/US00/05841
 PRIOR FILING DATE: March 2, 2000
 PRIOR APPLICATION NUMBER: PCT/US00/08439
 PRIOR FILING DATE: March 30, 2000
 PRIOR APPLICATION NUMBER: PCT/US00/14042
 PRIOR FILING DATE: May 22, 2000
 PRIOR APPLICATION NUMBER: PCT/US00/20710
 PRIOR FILING DATE: July 28, 2000
 PRIOR APPLICATION NUMBER: PCT/US00/32678
 PRIOR FILING DATE: December 1, 2000
 PRIOR APPLICATION NUMBER: PCT/US01/06520
 PRIOR FILING DATE: February 28, 2001
 NUMBER OF SEQ ID NOS: 120
 SEQ ID NO 42
 LENGTH: 243
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-09-944-413-42

Query Match 100.0%; Score 1325; DB 9; Length 243;
 Best Local Similarity 100.0%; Pred. No. 2, 3e-89;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGHHGSGGLPGRDGRGRCGAPG 60
 |||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGHHGSGGLPGRDGRGRCGAPG 60
 QY 61 APGEKGEGRGLGPGPDGPGCEAGPAGPTGAGCSVPSPASAKRSERVPSPD 120
 Db 61 APGEKGEGRGLGPGPDGPGCEAGPAGPTGAGCSVPSPASAKRSERVPSPD 120
 QY 121 ALPFDRLVNEQGHYAVTGKFTCOVPGVYFAVAATYRASLOPDLVKNESIASFFQ 180
 Db 121 ALPFDRLVNEQGHYAVTGKFTCOVPGVYFAVAATYRASLOPDLVKNESIASFFQ 180
 QY 181 FFGGMPKPSLSCGAMRLEPEDQVWVGVGDYIGIVASITDSTFSGFLVYSMHSSP 240
 Db 181 FFGGMPKPSLSCGAMRLEPEDQVWVGVGDYIGIVASITDSTFSGFLVYSMHSSP 240
 QY 241 VFA 243
 Db 241 VFA 243

RESULT 2
 US-09-944-403-42
 Sequence 42, Application US/09944403
 Patent No. US20020165143A1
 GENERAL INFORMATION:
 APPLICANT: Baker, Kevin
 APPLICANT: Bolstein, David
 APPLICANT: Baton, Dan
 APPLICANT: Ferrara, Napoleone
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gottsen, Mary
 APPLICANT: Goddard, Audrey
 APPLICANT: Grimaldi, Christopher
 APPLICANT: Gurney, Austin
 APPLICANT: Hillan, Kenneth
 APPLICANT: Kjaevn, Ivar
 APPLICANT: Napier, Mary
 APPLICANT: Roy, Margaret
 APPLICANT: Tumas, Daniel
 APPLICANT: Wood, William
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P2548P1C1
 CURRENT APPLICATION NUMBER: US/09/944,403
 PRIOR FILING DATE: 2001-09-26
 PRIOR APPLICATION NUMBER: 09/866,028
 PRIOR FILING DATE: 2001-05-25
 PRIOR APPLICATION NUMBER: 60/067,411
 PRIOR FILING DATE: December 3, 1997
 PRIOR APPLICATION NUMBER: 60/069,334
 PRIOR FILING DATE: December 11, 1997
 PRIOR APPLICATION NUMBER: 60/069,335
 PRIOR FILING DATE: December 11, 1997
 PRIOR APPLICATION NUMBER: 60/069,278
 PRIOR FILING DATE: December 11, 1997
 PRIOR APPLICATION NUMBER: 60/069,425
 PRIOR FILING DATE: December 12, 1997
 PRIOR APPLICATION NUMBER: 60/069,696
 PRIOR FILING DATE: December 16, 1997
 PRIOR APPLICATION NUMBER: 60/069,694
 PRIOR FILING DATE: December 16, 1997
 PRIOR APPLICATION NUMBER: 60/069,702
 PRIOR FILING DATE: December 16, 1997
 PRIOR APPLICATION NUMBER: 60/069,870
 PRIOR FILING DATE: December 17, 1997
 PRIOR APPLICATION NUMBER: 60/069,873
 PRIOR FILING DATE: December 17, 1997
 PRIOR APPLICATION NUMBER: 60/068,017
 PRIOR FILING DATE: December 18, 1997
 PRIOR APPLICATION NUMBER: 60/070,440
 PRIOR FILING DATE: January 5, 1998
 PRIOR APPLICATION NUMBER: 60/074,086
 PRIOR FILING DATE: February 9, 1998

```

PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020165143A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020165143A1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-403-42

```

```

Query Match 100.0% Score 1325 DB 9 Length 243
Best Local Similarity 100.0% Pred. No. 2.3e-89
Matches 243: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

1 MRPLVLLLLGLAASPLDNNKIPSLCPHGLPGTFFGHHGSGCLGRCGRDGRDAPC 60
1 MRPLVLLLLGLAASPLDNNKIPSLCPHGLPGTFFGHHGSGCLGRCGRDGRDAPC 60
61 APGEHGEGRGRLPGRCGPRGEGPAGTGTGAGCCSVPPRSARSAKRSERVPPSPD 120
61 APGEHGEGRGRLPGRCGPRGEGPAGTGTGAGCCSVPPRSARSAKRSERVPPSPD 120
121 APLEFDRVLVNEOGHYAVNGKFTCOVPGYVYFAVNAVTVRASIQEDLVNKGESIASFQ 180
121 APLEFDRVLVNEOGHYAVNGKFTCOVPGYVYFAVNAVTVRASIQEDLVNKGESIASFQ 180
181 FFGGMPRPASISGAGMVRLEPEDQWVQVGVGDIYIGTASIKTDSFSGFLVYSDMHSSP 240
181 FFGGMPRPASISGAGMVRLEPEDQWVQVGVGDIYIGTASIKTDSFSGFLVYSDMHSSP 240

```

```

DB 181 FFGGMPRPASISGAGMVRLEPEDQWVQVGVGDIYIGTASIKTDSFSGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

```

```

RESULT 3
US-09-944-896-42
Sequence 42, Application US/09944896
Patent No. US20020168715A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Batton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerltzen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tamas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998

```

```

PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020168715a1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020168715a1ember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03565
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06530
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-896-42

```

```

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pct. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNRKIPSLCPHPLPGTPRHSGSGCLPRGRGRGRCAGC 60
DB 1 MRPLVLLLLGLAAGSPPLDNRKIPSLCPHPLPGTPRHSGSGCLPRGRGRGRCAGC 60
QY 61 APGEKGGGRGRLPGPRGDPGRKAGPAGTPAGECVPRPSAFSAKRSRVPSPD 120
DB 61 APGEKGGGRGRLPGPRGDPGRKAGPAGTPAGECVPRPSAFSAKRSRVPSPD 120
QY 121 APLPFDVILVNEQGHDAVTKFTCOVPGYVFAVATYRASLQFDLVKNCESTASFQ 180
DB 121 APLPFDVILVNEQGHDAVTKFTCOVPGYVFAVATYRASLQFDLVKNCESTASFQ 180
QY 181 FFGGMPRPSALSGAWRLPEDOVWVGVYDITGIVASIKTDSFSGFLYSPWHSPP 240
DB 181 FFGGMPRPSALSGAWRLPEDOVWVGVYDITGIVASIKTDSFSGFLYSPWHSPP 240
QY 241 VFA 243
DB 241 VFA 243

```

```

RESULT 4
US-09-944-944-42
Sequence 42, Application US/09944944
Patent No. US20020173463A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David

```

```

APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Geritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kljavin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1c1
CURRENT APPLICATION NUMBER: US/09/944,944
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067,411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,666
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020173463A1ember 30, 1999

```

```

; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: No. US20020173463A1ember 30, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/28301
; PRIOR FILING DATE: December 1, 1999
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: December 16, 1999
; PRIOR APPLICATION NUMBER: PCT/US00/03565
; PRIOR FILING DATE: February 11, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: February 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: March 2, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/08439
; PRIOR FILING DATE: March 30, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: May 22, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/20710
; PRIOR FILING DATE: July 28, 2000
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: December 1, 2000
; PRIOR APPLICATION NUMBER: PCT/US01/006520
; PRIOR FILING DATE: February 28, 2001
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Saplen
US-09-944-944-42

Query Match      100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLGLAGSPPLDDNKIPSLCPGHGRLPCTPGHHSQGLPGRDGRDGRGAPG 60
DB 1 MRPLVLLGLAGSPPLDDNKIPSLCPGHGRLPCTPGHHSQGLPGRDGRDGRGAPG 60
QY 61 APGEKGEGRPGLPGRGDPGPRGEAGPAGPTGPAGECSVPPRSASFSAKRSRVPSPSD 120
DB 61 APGEKGEGRPGLPGRGDPGPRGEAGPAGPTGPAGECSVPPRSASFSAKRSRVPSPSD 120
QY 121 APLEPRDLVNEQGHYDAVTGKFTCOVPGVYFFAVHATVYRASLOFDLVKNGESIASFQ 180
DB 121 APLEPRDLVNEQGHYDAVTGKFTCOVPGVYFFAVHATVYRASLOFDLVKNGESIASFQ 180
QY 181 FEGGMPKPAISLGGAMVRLPEPEDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
DB 181 FEGGMPKPAISLGGAMVRLPEPEDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 5
US-09-944-907-42
; Sequence 42, Application US/09944907
; Publication No. US20020198147A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerlitsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Hillan, Kenneth
; APPLICANT: Kijavlin, Ivar
; APPLICANT: Napier, Mary
; APPLICANT: Roy, Margaret
```

```

; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/944,907
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Saplen
US-09-944-907-42

Query Match      100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLGLAGSPPLDDNKIPSLCPGHGRLPCTPGHHSQGLPGRDGRDGRGAPG 60
DB 1 MRPLVLLGLAGSPPLDDNKIPSLCPGHGRLPCTPGHHSQGLPGRDGRDGRGAPG 60
QY 61 APGEKGEGRPGLPGRGDPGPRGEAGPAGPTGPAGECSVPPRSASFSAKRSRVPSPSD 120
DB 61 APGEKGEGRPGLPGRGDPGPRGEAGPAGPTGPAGECSVPPRSASFSAKRSRVPSPSD 120
QY 121 APLEPRDLVNEQGHYDAVTGKFTCOVPGVYFFAVHATVYRASLOFDLVKNGESIASFQ 180
DB 121 APLEPRDLVNEQGHYDAVTGKFTCOVPGVYFFAVHATVYRASLOFDLVKNGESIASFQ 180
QY 181 FEGGMPKPAISLGGAMVRLPEPEDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
DB 181 FEGGMPKPAISLGGAMVRLPEPEDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 6
US-09-944-929-42
; Sequence 42, Application US/09944929
; Publication No. US20020197612A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerlitsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Hillan, Kenneth
; APPLICANT: Kijavlin, Ivar
; APPLICANT: Napier, Mary
; APPLICANT: Roy, Margaret
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/944,929
; CURRENT FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: 09/866,028
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
; LENGTH: 243
; TYPE: PRT
```


ORGANISM: Homo Sapien
US-09-944-929-42

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2,3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLGLGAGSPPLDDNKIPSLCPGHGCLPCTPHHHSQGLPGRDGHDGARG 60
DB 1 MRPLVLLGLGAGSPPLDDNKIPSLCPGHGCLPCTPHHHSQGLPGRDGHDGARG 60
QY 61 ABEKGEGRPGPLPGRBGRPGRGEAGPAGTPGAGECVSPRPSAFSAKRSRVPSPSD 120
DB 61 ABEKGEGRPGPLPGRBGRPGRGEAGPAGTPGAGECVSPRPSAFSAKRSRVPSPSD 120
QY 121 APLPDRVLYNQGHDAVTGKFTCOVPCVYFPAVHATVYRASLQFDLVKNGESIASFEQ 180
DB 121 APLPDRVLYNQGHDAVTGKFTCOVPCVYFPAVHATVYRASLQFDLVKNGESIASFEQ 180
QY 181 FFGGMPKPSLSCGAMVRLEPEDQYVWGVGDYGTASIKTDSFGFLVYSMDHSSP 240
DB 181 FFGGMPKPSLSCGAMVRLEPEDQYVWGVGDYGTASIKTDSFGFLVYSMDHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 7
US-10-028-072-362
Sequence 362, Application US/10028072
Publication No. US20030004311A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Mei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OF INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028,072
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059134
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588

PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079663
PRIOR FILING DATE: 1998-02-27
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/080165
PRIOR FILING DATE: 1998-03-31
PRIOR APPLICATION NUMBER: 60/081203
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081229
PRIOR FILING DATE: 1998-04-09
PRIOR APPLICATION NUMBER: 60/081695
PRIOR FILING DATE: 1998-04-14
PRIOR APPLICATION NUMBER: 60/081817
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081818
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082999
PRIOR FILING DATE: 1998-04-24
PRIOR APPLICATION NUMBER: 60/083322
PRIOR FILING DATE: 1998-04-28
PRIOR APPLICATION NUMBER: 60/083545
PRIOR FILING DATE: 1998-04-29
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084637
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085149
PRIOR FILING DATE: 1998-05-12
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086414
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/086430
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/087106
PRIOR FILING DATE: 1998-05-28
PRIOR APPLICATION NUMBER: 60/088026
PRIOR FILING DATE: 1998-06-04
PRIOR APPLICATION NUMBER: 60/088730
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088741
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088810
PRIOR FILING DATE: 1998-06-10
PRIOR APPLICATION NUMBER: 60/088858
PRIOR FILING DATE: 1998-06-11
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089599
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089907
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/089947
PRIOR FILING DATE: 1998-06-19
PRIOR APPLICATION NUMBER: 60/090349
PRIOR FILING DATE: 1998-06-23
PRIOR APPLICATION NUMBER: 60/090429
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090445

PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090538
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090863
PRIOR FILING DATE: 1998-06-26
PRIOR APPLICATION NUMBER: 60/091360
PRIOR FILING DATE: 1998-07-01
PRIOR APPLICATION NUMBER: 60/091519
PRIOR FILING DATE: 1998-07-02
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2,3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAASPPLDNKKITSLCPGCHGLPCTPCGHHSSQGLPGDGRDGDGARG 60
Db 1 MRPLVLLLLGLAASPPLDNKKITSLCPGCHGLPCTPCGHHSSQGLPGDGRDGDGARG 60
QY 61 APGEKGGGRPGLPGRGDPGRGEGAGPAGTPAGGECSSVPRSAFSAKRSESRVPPSPD 120
Db 61 APGEKGGGRPGLPGRGDPGRGEGAGPAGTPAGGECSSVPRSAFSAKRSESRVPPSPD 120
QY 121 ALPFDRLVNLNQGHRDAVTGKFTQVPGVYFAVAHATYRASLQFDLVKNCESTIASFFQ 180
Db 121 ALPFDRLVNLNQGHRDAVTGKFTQVPGVYFAVAHATYRASLQFDLVKNCESTIASFFQ 180
QY 181 FFGGMPKPSLSGSMVRLPEPDQYVWGVGDYIGIYASIKTDSFGGLVYSWMHSP 240
Db 181 FFGGMPKPSLSGSMVRLPEPDQYVWGVGDYIGIYASIKTDSFGGLVYSWMHSP 240
QY 241 VFA 243
Db 241 VFA 243

RESULT 8
US-10-121-049-362
Sequence 362, Application US/10121049
Publication No. US2003002239A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C17
CURRENT FILING DATE: 2002-04-12
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-121-049-362

Query Match 100.0%; Score 1325; DB 9; Length 243;

Best Local Similarity 100.0%; Pred. No. 2,3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPCTGCHHSGSLPGRDGRDCAAG 60
DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPCTGCHHSGSLPGRDGRDCAAG 60
OY 61 APGEKGGRRPGLPGRDPPGRGAGPAGTGPAGECVPPRSASFSAKRSSESVPPSPD 120
DB 61 APGEKGGRRPGLPGRDPPGRGAGPAGTGPAGECVPPRSASFSAKRSSESVPPSPD 120
OY 121 APLPFDRLVNVNCGHYDAVTGKFTCOVPGVYFFAVHATVYRASLQFDLVNKGESIASFFQ 180
DB 121 APLPFDRLVNVNCGHYDAVTGKFTCOVPGVYFFAVHATVYRASLQFDLVNKGESIASFFQ 180
OY 181 FEGGMPKPRASISGAMVRLPEDDVWVGVGDYIGIYASIKTDSSTFSGLVYSMDHSSP 240
DB 181 FEGGMPKPRASISGAMVRLPEDDVWVGVGDYIGIYASIKTDSSTFSGLVYSMDHSSP 240
OY 241 VFA 243
DB 241 VFA 243

RESULT 9

US-10-123-904-362
Sequence 362, Application US/10123904
Publication No. US20030022328A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C54
CURRENT APPLICATION NUMBER: US/10/123,904
CURRENT FILING DATE: 2002-04-16
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-123-904-362

Query Match
Best Local Similarity 100.0%; Score 1325; DB 9; Length 243;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPCTGCHHSGSLPGRDGRDCAAG 60
DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPCTGCHHSGSLPGRDGRDCAAG 60
OY 61 APGEKGGRRPGLPGRDPPGRGAGPAGTGPAGECVPPRSASFSAKRSSESVPPSPD 120
DB 61 APGEKGGRRPGLPGRDPPGRGAGPAGTGPAGECVPPRSASFSAKRSSESVPPSPD 120
OY 121 APLPFDRLVNVNCGHYDAVTGKFTCOVPGVYFFAVHATVYRASLQFDLVNKGESIASFFQ 180
DB 121 APLPFDRLVNVNCGHYDAVTGKFTCOVPGVYFFAVHATVYRASLQFDLVNKGESIASFFQ 180

DB

121 APLPFDRLVNVNCGHYDAVTGKFTCOVPGVYFFAVHATVYRASLQFDLVNKGESIASFFQ 180

OY

181 FEGGMPKPRASISGAMVRLPEDDVWVGVGDYIGIYASIKTDSSTFSGLVYSMDHSSP 240

DB

181 FEGGMPKPRASISGAMVRLPEDDVWVGVGDYIGIYASIKTDSSTFSGLVYSMDHSSP 240

OY

241 VFA 243

DB

241 VFA 243

RESULT 10

US-10-140-470-362
Sequence 362, Application US/10140470
Publication No. US20030022331A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerlitsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C160
CURRENT APPLICATION NUMBER: US/10/140,470
CURRENT FILING DATE: 2002-05-06
Prior Application removed - See File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-470-362

Query Match
Best Local Similarity 100.0%; Score 1325; DB 9; Length 243;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPCTGCHHSGSLPGRDGRDCAAG 60
DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPCTGCHHSGSLPGRDGRDCAAG 60
OY 61 APGEKGGRRPGLPGRDPPGRGAGPAGTGPAGECVPPRSASFSAKRSSESVPPSPD 120
DB 61 APGEKGGRRPGLPGRDPPGRGAGPAGTGPAGECVPPRSASFSAKRSSESVPPSPD 120
OY 121 APLPFDRLVNVNCGHYDAVTGKFTCOVPGVYFFAVHATVYRASLQFDLVNKGESIASFFQ 180
DB 121 APLPFDRLVNVNCGHYDAVTGKFTCOVPGVYFFAVHATVYRASLQFDLVNKGESIASFFQ 180
OY 181 FEGGMPKPRASISGAMVRLPEDDVWVGVGDYIGIYASIKTDSSTFSGLVYSMDHSSP 240
DB 181 FEGGMPKPRASISGAMVRLPEDDVWVGVGDYIGIYASIKTDSSTFSGLVYSMDHSSP 240
OY 241 VFA 243
DB 241 VFA 243

RESULT 11

US-09-796-753-68

QY	121	APLEPRVLYNCOGHADANTGKFTCCQPCVYVEAVHATVYRASLOFDLYKNGESIASFQ	180
Db	121	APLEPRVLYNCOGHADANTGKFTCCQPCVYVEAVHATVYRASLOFDLYKNGESIASFQ	180
QY	181	FEQGMKRPASLSGGAMVRLEPEDQVWVNGVCGDYITGYASTKTDSTSGFLVYSDMHS	240
Db	181	FEQGMKRPASLSGGAMVRLEPEDQVWVNGVCGDYITGYASTKTDSTSGFLVYSDMHS	240

OY 241 VFA 243
DB 241 VFA 243

RESULT 13

US-10-176-918-362
Sequence 362, Application US/10176918
Publication No. US20030027275A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Flivaotoff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C382
CURRENT APPLICATION NUMBER: US/10/176,918
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-918-362

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRPLVLLLLGLAGSPRLDDNKIPSLCPGHPGLPTGPHHGSQGLPRGDRGRGAGC 60
DB 1 MRPLVLLLLGLAGSPRLDDNKIPSLCPGHPGLPTGPHHGSQGLPRGDRGRGAGC 60
OY 61 APGEKGGGRGLPGPRGDRGPRGEGAPGPTGPAECVPPRSASFSAKRSRVPSPD 120
DB 61 APGEKGGGRGLPGPRGDRGPRGEGAPGPTGPAECVPPRSASFSAKRSRVPSPD 120
OY 121 APPLPDRVLNVEGCHYDAVTGKFTCOVPGVYFAVHATYRASLOPDLVKNESIASFFQ 180
DB 121 APPLPDRVLNVEGCHYDAVTGKFTCOVPGVYFAVHATYRASLOPDLVKNESIASFFQ 180
OY 181 FFGGMPKPSLSGAMVRLPEPDQVWVGVGYDYGITIASITDSTFSGFLVYSMDHSSP 240
DB 181 FFGGMPKPSLSGAMVRLPEPDQVWVGVGYDYGITIASITDSTFSGFLVYSMDHSSP 240
OY 241 VFA 243
DB 241 VFA 243

RESULT 14

US-10-176-921-362
Sequence 362, Application US/10176921
Publication No. US20030027276A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen

APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Flivaotoff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C288
CURRENT APPLICATION NUMBER: US/10/176,921
CURRENT FILING DATE: 2002-06-20
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-921-362

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRPLVLLLLGLAGSPRLDDNKIPSLCPGHPGLPTGPHHGSQGLPRGDRGRGAGC 60
DB 1 MRPLVLLLLGLAGSPRLDDNKIPSLCPGHPGLPTGPHHGSQGLPRGDRGRGAGC 60
OY 61 APGEKGGGRGLPGPRGDRGPRGEGAPGPTGPAECVPPRSASFSAKRSRVPSPD 120
DB 61 APGEKGGGRGLPGPRGDRGPRGEGAPGPTGPAECVPPRSASFSAKRSRVPSPD 120
OY 121 APPLPDRVLNVEGCHYDAVTGKFTCOVPGVYFAVHATYRASLOPDLVKNESIASFFQ 180
DB 121 APPLPDRVLNVEGCHYDAVTGKFTCOVPGVYFAVHATYRASLOPDLVKNESIASFFQ 180
OY 181 FFGGMPKPSLSGAMVRLPEPDQVWVGVGYDYGITIASITDSTFSGFLVYSMDHSSP 240
DB 181 FFGGMPKPSLSGAMVRLPEPDQVWVGVGYDYGITIASITDSTFSGFLVYSMDHSSP 240
OY 241 VFA 243
DB 241 VFA 243

RESULT 15

US-10-137-865-362
Sequence 362, Application US/10137865
Publication No. US20030032155A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Flivaotoff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K

APPLICANT: Wood,William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C154
CURRENT APPLICATION NUMBER: US/10/137,865
CURRENT FILING DATE: 2002-05-03
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-137-865-362

Query Match 100.0%; Score 1325; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 2.3e-89;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MRPLVLLLLGLAGSPPLDNNKIPSLCPGHPGLPCTPGHHSQGLPRGDRGDRGAPG 60
Db 1 MRPLVLLLLGLAGSPPLDNNKIPSLCPGHPGLPCTPGHHSQGLPRGDRGDRGAPG 60
QY 61 APGEKGGGRGLPGRGDPPRGEGAPAGPTGAGCEGVPRSAFSAKRSRVPSPD 120
Db 61 APGEKGGGRGLPGRGDPPRGEGAPAGPTGAGCEGVPRSAFSAKRSRVPSPD 120
QY 121 APLPDRVLNVEGCHDAVTGKFTCOVPGYVFAVHATYVRASLQFDLVKNGESIASFFQ 180
Db 121 APLPDRVLNVEGCHDAVTGKFTCOVPGYVFAVHATYVRASLQFDLVKNGESIASFFQ 180
QY 181 FFGGMPKPASLSGAVRLEPEDQVWVQVGVGDYIGIYASIKTDSFEFGFLVSDWHSSP 240
Db 181 FFGGMPKPASLSGAVRLEPEDQVWVQVGVGDYIGIYASIKTDSFEFGFLVSDWHSSP 240
QY 241 VFA 243
Db 241 VFA 243

Search completed: June 18, 2003, 15:02:35
Job time : 57 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 14:54:03 ; Search time 71 Seconds

(Without alignments)
456.055 Million cell updates/sec

Title: US-09-943-851A-42

Sequence: 1 MRPLVLVLLGLAAGSPPLD.....DSTFGSLVYSDMHSSVPFA 243

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size : 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

A_Geneseq_101002:*

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	243	100.0	243	20	AAV06481 Human tumour-assoc
2	243	100.0	243	20	AAV17827 Human PRO344 prote
3	243	100.0	243	20	AAW97984 Human adipocyte-sp
4	243	100.0	243	21	AAW33461 Human PRO344 prote
5	243	100.0	243	21	AAW71468 Human PRO344 prote
6	243	100.0	243	21	AAW93588 Human acid sequenc
7	243	100.0	243	21	AAW01318 Human PRO344 polyp
8	243	100.0	243	22	AAW12352 Human TANGO 253 SE
9	243	100.0	243	22	AAW65815 Human adipocyte co
10	243	100.0	243	22	AAW49593 Human adipocyte co

11	243	100.0	243	22	AAW49599 Human adipocyte co
12	228	93.8	228	22	AAW65816 Human mature TANGO
13	213	87.7	220	22	ABG12724 Novel human diago
14	185	76.1	243	22	AAW65888 Human secreted pro
15	174	71.6	243	22	AAW65891 Human secreted pro
16	173	71.2	201	22	AAO21663 Human secreted pro
17	148	60.9	243	22	AAW65889 Human secreted pro
18	142	58.4	243	22	AAW65890 Human secreted pro
19	128	52.7	128	22	AAW65819 Human TANGO 253 CI
20	82	33.7	151	22	ABG12723 Novel human diago
21	70	28.8	243	21	AAW76040 Rat skin cell prot
22	70	28.8	243	22	AAW55979 Skin cell protein,
23	70	28.8	243	22	ABW12179 Rat protein isolat
24	62	25.5	228	22	AAW65821 Murine mature TANG
25	62	25.5	242	22	AAW65896 Murine secreted pr
26	62	25.5	243	22	AAW65820 Murine TANGO 253 S
27	62	25.5	243	22	AAW65897 Murine secreted pr
28	62	25.5	243	22	AAW65898 Murine secreted pr
29	62	25.5	243	22	AAW65899 Murine secreted pr
30	60	24.7	60	22	AAW65818 Murine TANGO 253 C
31	51	21.0	128	22	AAW65824 Murine TANGO 253 C
32	43	17.7	60	22	AAW65823 Human signal pepti
33	30	12.3	87	21	AAW87269 Human TANGO 253 SI
34	15	6.2	15	22	AAW65817 Human TANGO 253 SI
35	11	4.5	27	22	AAW65817 Human TANGO 253 SI
36	11	4.5	56	22	AAW65817 Human TANGO 253 SI
37	11	4.5	56	22	AAW65817 Human TANGO 253 SI
38	11	4.5	56	22	AAW65817 Human TANGO 253 SI
39	11	4.5	56	22	AAW65817 Human TANGO 253 SI
40	11	4.5	56	22	AAW65817 Human TANGO 253 SI
41	11	4.5	56	22	AAW65817 Human TANGO 253 SI
42	11	4.5	56	22	AAW65817 Human TANGO 253 SI
43	11	4.5	56	22	AAW65817 Human TANGO 253 SI
44	11	4.5	56	23	ABG39329 Human peptide enco
45	11	4.5	128	20	AAV02690 Human secreted pro

ALIGNMENTS

RESULT 1				
ID	AAV06481	standard; protein; 243 AA.		
XX	AAV06481;			
AC	27-SEP-1999	(first entry)		
XX				
DT				
XX				
DE	Human tumour-associated protein PRO344.			
XX				
KW	PRO344; UNQ303; cancer; tumour; diagnosis; therapy; human.			
XX				
OS	Homo sapiens.			
XX				
PH	Key	Location/Qualifiers		
FT	Peptide	1..15		
FT		/note="signal peptide"		
FT	Protein	16..243		
FT		/note="mature protein"		
FT	Modified-site	68..215		
FT		/note="N-myristoylated"		
FT	Modified-site	216..243		
FT		/note="N-myristoylated"		
XX				
PN	MO9935170-AZ.			
XX				
XX	15-JUL-1999.			
PD				
XX				
XX	05-JAN-1999.	99WO-US00106.		
PF				
XX				
PR	20-NOV-1998.	98US-0109304.		
PR	05-JAN-1998.	98US-0070440.		
PR	29-APR-1998.	98US-0083500.		

PR 22-MAY-1998; 98US-0086414.
PR 10-JUN-1998; 98US-0088742.
PR 10-NOV-1998; 98US-0107783.

XX (GERTH) GENENTECH INC.

PI Botstein D, Goddard A, Gurney AL, Hillian KJ, Lawrence DA;
PI Roy MA, Wood WI;

DR WPI; 1999-430385/36.
DR N-PSDB; AAX87258.

PT Antibody against proteins expressed in neoplastic cells, useful for
PT tumor diagnosis and treatment

PS Example 1; Fig 10; 162pp; English.

CC This sequence represents human PRO344 (UNQ303), a protein encoded
CC by the novel cDNA clone DNA40592 (see AAX87258). Amplification of
CC DNA40592 was observed in primary lung tumours and in primary colon
CC tumours, suggesting a significant role in tumour formation and
CC growth. Antagonists (e.g. antibodies) directed to PRO344 may have
CC use in cancer therapy. The invention identifies 14 genes (see
CC AAX87254-67) that are amplified in the genome of tumour cells. Such
CC amplification is expected to be associated with overexpression of
CC the gene product and to contribute to tumorigenesis. The encoded
CC proteins (see AAY06477-90) may be useful targets for the diagnosis
CC and/or treatment (including prevention) of certain cancers, and may
CC act as predictors of the prognosis of tumour treatment. Antibodies
CC that bind the proteins are claimed and used in claimed cancer
CC diagnostic kits.

XX Sequence 243 AA:

Query Match 100.0%; Score 243; DB 20; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.9e-212;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGTHHSGGLPGRDGRDGAAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGTHHSGGLPGRDGRDGAAPG 60
QY 61 APGEKGEGRGRLPGPRDPCPRGEGAPAGTGPAGCSVPFSAKSESSESVPPSD 120
DB 61 APGEKGEGRGRLPGPRDPCPRGEGAPAGTGPAGCSVPFSAKSESSESVPPSD 120
QY 121 APLEFDRVLVNEQGHYDAVTKFTCOVPGYFFAVHATYRASLQFDLVKNGESIASFQ 180
DB 121 APLEFDRVLVNEQGHYDAVTKFTCOVPGYFFAVHATYRASLQFDLVKNGESIASFQ 180
QY 181 FFGGMPKPRASLSGAMVRLPEPDQVWVQVGVGDYIGIYASIKTDSFSGFLVYSMDHSSP 240
DB 181 FFGGMPKPRASLSGAMVRLPEPDQVWVQVGVGDYIGIYASIKTDSFSGFLVYSMDHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 2

AAI17827 standard; Protein; 243 AA.

AAI17827;

XX 12-AUG-1999 (first entry)

DE Human PRO344 protein sequence.

KW Human: PRO protein; tumour necrosis factor family; TNF; cytokine;
KM secreted protein; transmembrane protein; inflammation disorder.
XX Homo sapiens.

XX MO9928462-A2.
XX 10-JUN-1999.

XX 01-DEC-1998; 98WO-US25108.

XX 25-FEB-1998; 98US-0075945.

XX 03-DEC-1997; 97US-0067411.

XX 11-DEC-1997; 97US-0069278.

XX 11-DEC-1997; 97US-0069334.

XX 11-DEC-1997; 97US-0069335.

XX 12-DEC-1997; 97US-0069425.

XX 16-DEC-1997; 97US-0069694.

XX 16-DEC-1997; 97US-0069702.

XX 17-DEC-1997; 97US-0069870.

XX 18-DEC-1997; 97US-0069873.

XX 05-JAN-1998; 97US-0068017.

XX 09-FEB-1998; 98US-0070440.

XX 09-FEB-1998; 98US-0074086.

XX (GERTH) GENENTECH INC.

XX Baker KP, Chen J, Goddard A, Gurney AL, Wood WI;

XX WPI; 1999-371118/31.

XX N-PSDB; AAX80052.

XX Nucleic acids encoding PRO secreted and transmembrane proteins

XX Claim 12; Fig 21; 123pp; English.

XX The present invention describes nucleic acids encoding PRO secreted and

XX transmembrane proteins used therapeutically. The PRO proteins have

XX cytotactic, anti-inflammatory, anti-proliferative and immunosuppressive

XX activity. The proteins and polynucleotides can be used in therapy,

XX identification of homologues, raising antibodies and design of probes

XX and primers. They can be used in a range of diseases related to proteins

XX that they have homology with, e.g. a PRO protein having homology to

XX complement proteins may be used in inflammatory responses.

XX Sequence 243 AA:

Query Match 100.0%; Score 243; DB 20; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.9e-212;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGTHHSGGLPGRDGRDGAAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCPGHPGLPTGTHHSGGLPGRDGRDGAAPG 60
QY 61 APGEKGEGRGRLPGPRDPCPRGEGAPAGTGPAGCSVPFSAKSESSESVPPSD 120
DB 61 APGEKGEGRGRLPGPRDPCPRGEGAPAGTGPAGCSVPFSAKSESSESVPPSD 120
QY 121 APLEFDRVLVNEQGHYDAVTKFTCOVPGYFFAVHATYRASLQFDLVKNGESIASFQ 180
DB 121 APLEFDRVLVNEQGHYDAVTKFTCOVPGYFFAVHATYRASLQFDLVKNGESIASFQ 180
QY 181 FFGGMPKPRASLSGAMVRLPEPDQVWVQVGVGDYIGIYASIKTDSFSGFLVYSMDHSSP 240
DB 181 FFGGMPKPRASLSGAMVRLPEPDQVWVQVGVGDYIGIYASIKTDSFSGFLVYSMDHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 3

AAW97984

ID	AAW97984	standard: Protein: 243 AA.
XX	AC	AAW97984;
XX	DT	21-JUN-1999 (first entry)
XX	DE	Human adipocyte-specific protein zs1939.
XX	KM	Adipocyte-specific protein: zs1939; human; fatty acid metabolism; energy balance; nutrition; antimicrobial.
XX	OS	Homo sapiens.
XX	XX	
XX	Key	Location/Qualifiers
XX	Peptide	1..15
XX	FT	/note="signal peptide, alternatively the signal
XX	FT	peptide comprises residues 1..18"
XX	FT	16..243
XX	FT	/note="mature protein, alternatively the mature
XX	FT	protein comprises residues 19..243
XX	FT	(specifically claimed in Claim 4)"
XX	FT	30..96
XX	FT	/note="collagen-like domain"
XX	FT	98..243
XX	FT	/note="globular domain"
XX	FT	105..109
XX	FT	/note="beta strand"
XX	FT	128..130
XX	FT	/note="beta strand"
XX	FT	136..139
XX	FT	/note="beta strand"
XX	FT	143..146
XX	FT	/note="beta strand"
XX	FT	164..171
XX	FT	/note="beta strand"
XX	FT	176..182
XX	FT	/note="beta strand"
XX	FT	187..200
XX	FT	/note="beta strand"
XX	FT	204..210
XX	FT	/note="beta strand"
XX	FT	226..231
XX	FT	/note="beta strand"
XX	FT	111..135
XX	FT	/note="receptor binding domain"
XX	FT	170..174
XX	FT	/note="receptor binding domain"
XX	PN	WO9910492-A1.
XX	PD	04-MAR-1999.
XX	PF	26-AUG-1998; 98WO-US17724.
XX	PR	26-AUG-1997; 97US-0056983.
XX	XX	
XX	PA	(Zymo) ZYMOGENETICS INC.
XX	XX	
XX	PI	Humes JM, Sheppard PO;
XX	XX	
XX	DR	WPI; 1999-204665/17.
XX	DR	N-PSDB; AAX24684.
XX	XX	
XX	FT	zs1939 protein - used to modulate fatty acid metabolism
XX	PS	Claim 1; 111-112; 132pp; English.

This polypeptide comprises human adipocyte-specific protein zs1939, a protein that modulates free fatty acid metabolism. zs1939 is a member of a family of proteins having a globular domain and a collagen-like domain capable of dimerisation or oligomerisation. zs1939 polypeptides were initially identified by querying an EST database for secretory signal sequences characterised by an upstream

[illegible]

XX 14-SEP-2000.
 PD 02-MAR-2000; 2000MO-US05841.
 XX 08-MAR-1999; 99MO-US05028.
 XX 10-MAR-1999; 99US-0123618.
 PR 12-MAR-1999; 99US-0123957.
 PR 23-MAR-1999; 99US-0125775.
 PR 12-APR-1999; 99US-0128849.
 PR 20-APR-1999; 99MO-US08615.
 PR 28-APR-1999; 99US-0131445.
 PR 04-MAY-1999; 99US-0132371.
 PR 14-MAY-1999; 99US-0134287.
 PR 02-JUN-1999; 99MO-US12252.
 PR 23-JUN-1999; 99US-0141037.
 PR 20-JUL-1999; 99US-0144758.
 PR 26-JUL-1999; 99US-0145698.
 PR 28-JUL-1999; 99US-0146222.
 PR 01-SEP-1999; 99MO-US20511.
 PR 08-SEP-1999; 99MO-US20594.
 PR 13-SEP-1999; 99MO-US20944.
 PR 15-SEP-1999; 99MO-US21090.
 PR 15-SEP-1999; 99MO-US21547.
 PR 05-OCT-1999; 99MO-US23089.
 PR 29-OCT-1999; 99US-0162506.
 PR 29-NOV-1999; 99MO-US28214.
 PR 30-NOV-1999; 99MO-US28313.
 PR 30-NOV-1999; 99MO-US28409.
 PR 01-DEC-1999; 99MO-US28301.
 PR 01-DEC-1999; 99MO-US28634.
 PR 02-DEC-1999; 99MO-US28551.
 PR 02-DEC-1999; 99MO-US28564.
 PR 16-DEC-1999; 99MO-US28565.
 PR 30-DEC-1999; 99MO-US30095.
 PR 30-DEC-1999; 99MO-US30999.
 PR 05-JAN-2000; 2000MO-US00219.
 PR 06-JAN-2000; 2000MO-US00277.
 PR 06-JAN-2000; 2000MO-US00376.
 PR 11-FEB-2000; 2000MO-US03565.
 PR 18-FEB-2000; 2000MO-US04341.
 PR 18-FEB-2000; 2000MO-US04342.
 PR 22-FEB-2000; 2000MO-US04414.
 XX (GETH) GENENTECH INC.
 PA Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W,
 XX Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V,
 PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M,
 PI WPI; 2000-572271/53.
 DR N-PSDB; MAC58626.
 XX
 PT Sixty four PRO polypeptides, useful in the diagnosis and treatment of
 PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus -
 PS
 PS Claim 33; Fig 96; 309pp; English.

CC immunological diseases of the lung, and transplantation associated
 CC diseases including graft rejection and graft-versus-host disease.
 CC AAC58397 to AAC58578 represent PCR primers and hybridisation probes used
 CC in the isolation of human PRO sequences. AAC58579 to AAC58642 and
 CC AAB33414 to AAB33477 represent human PRO polynucleotide and protein
 CC sequences given in the exemplification of the present invention.
 XX
 SQ Sequence 243 AA:
 Query Match 100.0%; Score 243; DB 21; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4.9e-212;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPTGPHHSGQLGPRGRDGRDGAAPG 60
 Db 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPTGPHHSGQLGPRGRDGRDGAAPG 60
 QY 61 APGEKGGGRGRLPGPRDGPGEAGPACPTGAGCSVPPPSAFSAKSESRRVPPSD 120
 Db 61 APGEKGGGRGRLPGPRDGPGEAGPACPTGAGCSVPPPSAFSAKSESRRVPPSD 120
 QY 121 APLEFDRVLVNEQGHYDAVTGKFTCOYGVYFAVHATVYRASLQFDLVKNGESTIASFPQ 180
 Db 121 APLEFDRVLVNEQGHYDAVTGKFTCOYGVYFAVHATVYRASLQFDLVKNGESTIASFPQ 180
 QY 181 FFGGMPKPAASLSGAWRLEPEQVWVGVGYDITIVASIKTDSFSGFLVYSDMHSSP 240
 Db 181 FFGGMPKPAASLSGAWRLEPEQVWVGVGYDITIVASIKTDSFSGFLVYSDMHSSP 240
 QY 241 VFA 243
 Db 241 VFA 243
 RESULT 5
 AAY71468
 ID AAY71468 standard; Protein; 243 AA.
 XX
 AC AAY71468;
 XX
 DT 08-NOV-2000 (first entry)
 XX
 DE Human PRO344 protein.
 XX
 KW PRO344; DNA40592-1242; human; ATCC No: 209492; antiproliferative;
 KW neoplastic cell growth inhibitor; cytostatic; treatment; cancer; tumour;
 KW breast; prostate; colon; lung; renal; ovarian; central nervous system;
 KW CNS; leukemia; melanoma; Expressed Sequence Tag; EST; secreted protein;
 KW extracellular domain; ECD.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..15
 FT /Label= Signal_peptide
 FT Modified-site 11..17
 FT /note= "N-myristoylation site"
 FT Protein 16..243
 FT /Label= Mature_PRO344_protein
 FT Modified-site 68..74
 FT /note= "N-myristoylation site"
 FT Binding-site 77..80
 FT /note= "Cell attachment sequence"
 FT Modified-site 216..222
 FT /note= "N-myristoylation site"
 PD MO200032778-A2.
 PD 08-JUN-2000.
 XX 30-NOV-1999; 99MO-US28409.
 XX 01-DEC-1998; 98MO-US25108.

PR 16-DEC-1998: 98US-0112850.
 PR 22-DEC-1998: 98US-0113296.
 PR 20-JUL-1999: 99US-0144758.
 PR 26-JUL-1999: 99US-0145698.

XX (GENTH) GENENTECH INC.

PI Chen J, Goddard A, Gurney AL, Hillan K, Napier M, Wood WI;

XX MPI: 2000-412325/35.

DR N-PSDB: AAD01241.

XX New composition useful for inhibiting neoplastic cell growth and for
 PT treating cancers, comprises PRO655, PRO344 or PRO364 polypeptide or
 PT their antagonists

PS Claim 31: Fig 6: 108pp; English.

XX The present sequence is the human PRO344 protein, encoded by the cDNA
 CC clone, designated as DNA40592-1242. It is isolated from human foetal
 CC lung tissue, cDNA library. Identified using probes based on a consensus
 CC sequence DNA44398, derived from secreted protein extracellular domain
 CC (ECD) expressed sequence tag (EST). This clone is assigned ATCC deposit
 CC No: 209492. PRO344 functions as a neoplastic cell growth inhibitor and
 CC is used for treating tumours, using an effective amount of PRO655, PRO364
 CC and PRO344. This composition is especially useful for treatment of human
 CC cancers such as breast, prostate, colon, lung, renal, ovarian and CNS,
 CC leukemia and melanoma.

XX Sequence 243 AA:

Query Match 100.0%; Score 243; DB 21; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4.9e-212.

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRRLVLLLLGLAAGSPRLDNNKIPSLCPGHPGLPCTPGHNSGGLPGRGDRGDRGAPG 60

DB 1 MRRLVLLLLGLAAGSPRLDNNKIPSLCPGHPGLPCTPGHNSGGLPGRGDRGDRGAPG 60

QY 61 APGEKSGRPGRLPGPRGDPGRGEGAPRGPRGACSVPPRSASAKRSERVRPPSD 120

DB 61 APGEKSGRPGRLPGPRGDPGRGEGAPRGPRGACSVPPRSASAKRSERVRPPSD 120

QY 121 APPLPFDRLVNEGGHYDAVTGKFTCPGVYFAVAHATYRASLQFDLVKNGESIASFPQ 180

DB 121 APPLPFDRLVNEGGHYDAVTGKFTCPGVYFAVAHATYRASLQFDLVKNGESIASFPQ 180

QY 181 FFGGWRPASTLSCGAMVRLEPEQVWVQGVGDIYIGIVASIKTDSFSGFLVYSDMHSSP 240

DB 181 FFGGWRPASTLSCGAMVRLEPEQVWVQGVGDIYIGIVASIKTDSFSGFLVYSDMHSSP 240

QY 241 VFA 243

DB 241 VFA 243

QY 241 VFA 243

DB 241 VFA 243

QY 241 VFA 243

DB 241 VFA 243

QY 241 VFA 243

DB 241 VFA 243

QY 241 VFA 243

DB 241 VFA 243

QY 241 VFA 243

DB 241 VFA 243

QY 241 VFA 243

DB 241 VFA 243

QY 241 VFA 243

DB 241 VFA 243

QY 241 VFA 243

FT Peptide 1..15
 FT /note- "signal peptide"
 FT Modified-site 11..17
 FT /note- "N-myristoylation site"
 FT Modified-site 68..74
 FT /note- "N-myristoylation site"
 FT Region 77..80
 FT /note- "cell attachment sequence"
 FT Modified-site 216..222
 FT /note- "N-myristoylation site"

XX WO200037640-A2.

XX 29-JUN-2000.

XX 16-DEC-1999: 99WO-US30095.

XX 22-DEC-1998: 98US-0113296.

XX 08-MAR-1999: 99WO-US05028.

XX 02-JUN-1999: 99WO-US12252.

XX 01-SEP-1999: 99WO-US20111.

XX 15-SEP-1999: 99WO-US21090.

XX 30-NOV-1999: 99WO-US28313.

XX 01-DEC-1999: 99WO-US28409.

XX 02-DEC-1999: 99WO-US28301.

XX 02-DEC-1999: 99WO-US28565.

XX (GENTH) GENENTECH INC.

XX Botstein D, Goddard A, Gurney AL, Hillan K, Lawrence DA, Roy MA;

XX Wood WI;

XX MPI: 2000-452188/39.

XX N-PSDB: AAA46907.

XX New anti-polypeptide antibody useful in the treatment and diagnosis of

XX neoplastic cell growth and proliferation

XX Claim 61: Fig 10: 220pp; English.

XX The present sequence represents a novel human polypeptide. The

XX specification describes novel polypeptides designated PRO201, PRO292,

XX PRO327, PRO1265, PRO344, PRO347, PRO357, PRO1017,

XX PRO1112, PRO509, PRO853 and PRO882. These genes are amplified in

XX the genome of tumour cells. The polypeptides are believed to contribute

XX to tumorigenesis. The polypeptides are useful target for the

XX identification of certain cancers, and may act as predictors of the

XX prognosis of tumour treatment. Antibodies against these polypeptides

XX are useful in the treatment and diagnosis of neoplastic cell growth

XX and proliferation in mammals.

XX Sequence 243 AA:

QY 1 MRRLVLLLLGLAAGSPRLDNNKIPSLCPGHPGLPCTPGHNSGGLPGRGDRGDRGAPG 60

DB 1 MRRLVLLLLGLAAGSPRLDNNKIPSLCPGHPGLPCTPGHNSGGLPGRGDRGDRGAPG 60

QY 61 APGEKSGRPGRLPGPRGDPGRGEGAPRGPRGACSVPPRSASAKRSERVRPPSD 120

DB 61 APGEKSGRPGRLPGPRGDPGRGEGAPRGPRGACSVPPRSASAKRSERVRPPSD 120

QY 121 APPLPFDRLVNEGGHYDAVTGKFTCPGVYFAVAHATYRASLQFDLVKNGESIASFPQ 180

DB 121 APPLPFDRLVNEGGHYDAVTGKFTCPGVYFAVAHATYRASLQFDLVKNGESIASFPQ 180

QY 181 FFGGWRPASTLSCGAMVRLEPEQVWVQGVGDIYIGIVASIKTDSFSGFLVYSDMHSSP 240

DB 181 FFGGWRPASTLSCGAMVRLEPEQVWVQGVGDIYIGIVASIKTDSFSGFLVYSDMHSSP 240

OY 241 VFA 243
 DB 241 VFA 243

RESULT 7

AA01318 standard; Protein: 243 AA.

AA01318;

25-SEP-2000 (first entry)

Human PRO344 polypeptide.

PRO: Membrane bound protein; secreted protein: PRO357; PRO327;
 PRO343; PRO315; PRO241; PRO323; PRO299; PRO333; PRO344; PRO347;
 PRO353; PRO361; PRO365; transmembrane polypeptide;
 antibody; screening; detection; inhibition; probe; primer; human.

Homo sapiens.

Key Location/Qualifiers

Peptide 1..15

Modified-site /label= Signal peptide

Modified-site /note= "N-myristoylation site"

Modified-site /note= "N-myristoylation site"

Modified-site /label= Cell attachment sequence

Modified-site /note= "N-myristoylation site"

WO200032776-A2.

08-JUN-2000.

01-DEC-1999; 99WO-US28301.

01-DEC-1998; 98WO-US25108.

16-DEC-1998; 98US-0112850.

22-DEC-1998; 98US-0113296.

(GETH) GENENTECH INC.

Baker KP, Botstein D, Eaton DL, Ferrara N, Filvaroff E;

Gerritsen ME, Goddard A, Godowski PJ, Grimaldi CJ, Gurney AL;

Hillman KJ, Kljavin IJ, Napier MA, Roy MA, Tumas D, Wood WT;

WPI: 2000-412324/35.

N-PSDB: AAA49560.

New human nucleic acids encoding secreted and transmembrane

polypeptides, designated as PRO polypeptides, useful as pharmaceutical

and diagnostic agents

Claim 12; Fig 18; 187pp; English.

New human nucleic acids encoding secreted and transmembrane

polypeptides which are designated as PRO polypeptides are described

including as pharmaceutical and diagnostic agents. The membrane-bound

proteins can also be employed for screening of potential polypeptide or

small molecule inhibitors of the relevant receptor/ligand interaction.

Anti-PRO antibodies are useful for the affinity purification of PRO

from recombinant cell culture or natural sources.

Sequence 243 AA:

Query Match 100.0%; Score 243; DB 21; Length 243;

Best Local Similarity 100.0%; Pred No. 4, 9e-212;

Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHGSQGLPCRGDRGDRGAPG 60
 DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHGSQGLPCRGDRGDRGAPG 60
 OY 61 APGEKGGRRPGLPGRDPRGPRGAGPAGTGTAGGCSVPPRASFSAKRSSESVPPSD 120
 DB 61 APGEKGGRRPGLPGRDPRGPRGAGPAGTGTAGGCSVPPRASFSAKRSSESVPPSD 120
 OY 121 APLEPDRVLVNEQGHYDAVTGFTQVPGVYFFFAVHATVYRASLQDFLVNKGESIASFQ 180
 DB 121 APLEPDRVLVNEQGHYDAVTGFTQVPGVYFFFAVHATVYRASLQDFLVNKGESIASFQ 180
 OY 181 FFGGMPKPRASLSGAMVRLPEPDQVWQVGVGDYIGIYASIKTDSFSGFLVYSDMHSSP 240
 DB 181 FFGGMPKPRASLSGAMVRLPEPDQVWQVGVGDYIGIYASIKTDSFSGFLVYSDMHSSP 240
 OY 241 VFA 243
 DB 241 VFA 243

RESULT 8

AA012352 standard; Protein: 243 AA.

AA012352;

24-OCT-2001 (first entry)

Human PRO344 polypeptide sequence.

Human secretory and transmembrane; PRO: mammalian; cancer; lung;
 breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
 cartilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 adipocyte; A-peptide; factor VIIA; gene therapy.

Homo sapiens.

WO200140466-A2.

07-JUN-2001.

01-DEC-2000; 2000WO-US32678.

01-DEC-1999; 99WO-US28301.

01-DEC-1999; 99WO-US28634.

02-DEC-1999; 99WO-US28551.

02-DEC-1999; 99WO-US28564.

09-DEC-1999; 99US-0170262.

16-DEC-1999; 99WO-US30095.

20-DEC-1999; 99WO-US30911.

20-DEC-1999; 99WO-US30999.

30-DEC-1999; 99WO-US31243.

06-JAN-2000; 2000WO-US00277.

06-JAN-2000; 2000WO-US00376.

11-FEB-2000; 2000WO-US03565.

18-FEB-2000; 2000WO-US04341.

22-FEB-2000; 2000WO-US04342.

24-FEB-2000; 2000WO-US04914.

24-FEB-2000; 2000WO-US05004.

01-MAR-2000; 2000WO-US05601.

20-MAR-2000; 2000WO-US07377.

21-MAR-2000; 2000WO-US07532.

30-MAR-2000; 2000WO-US08439.

17-MAY-2000; 2000WO-US13705.

22-MAY-2000; 2000WO-US14042.

30-MAY-2000; 2000WO-US14941.

02-JUN-2000; 2000WO-US15264.

10-NOV-2000; 2000WO-US30873.

PA (GETH) GENENTECH INC.

XX Baker KP, Bersini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
XX Geritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
XX Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WL, Zhang Z;
XX WPI: 2001-408281/43.
XX N-PSDB: AAS21424.

PT Isolated, secretory and transmembrane PRO polypeptide used to detect
PT other PRO polypeptides, link bioactive molecules to cells expressing
PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
PT lung, breast, prostate, cervical

XX Claim 12: Fig 362; 813pp; English.

XX AAU2172-AAU2446 represent novel human secretory and transmembrane
XX PRO polypeptides. The PRO polypeptides are useful to detect other
XX PRO polypeptides, to link bioactive molecules to cells expressing
XX PRO polypeptides, to modulate biological activities of cells expressing
XX PRO polypeptides, and to detect the presence of mammalian lung, colon,
XX breast, prostate, rectal, cervical or liver tumours by comparing PRO
XX polypeptide expression in a cell sample to that in a control sample.
XX Some of the 275 sequences are also useful to stimulate the release of
XX tumour necrosis factor-alpha (TNF-alpha) from human blood, the
XX proliferation or differentiation of chondrocytes, the proliferation or
XX gene expression in pericyte cells, the release of proteoglycans from
XX cartilage, the proliferation of inner ear utricular supporting cells or
XX of T-lymphocytes, the release of a cytokine from peripheral blood
XX monocytes (PBMCs), or the proliferation of endothelial cells. Some of
XX the PRO polypeptides may modulate glucose or free fatty acid uptake by
XX skeletal muscle cells or by adipocytes; or inhibit binding of A-polypeptide
XX to factor VIIA. The PRO polypeptides can be used in assays to identify
XX molecules involved in binding interactions. The polynucleotides encoding
XX PRO polypeptides can be used to generate probes, antisense RNA/DNA,
XX transgenic or knock out animals and can be used in gene therapy.

XX Sequence 243 AA:

Query Match 100.0%; Score 243; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.9e-212;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTPGHNGSGLRGDRGDRGAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTPGHNGSGLRGDRGDRGAPG 60
QY 61 APGEKGEGRGRLPGRCRDPGRGEGAPGPTGPAECVSPPRSASAKRSSESVPPSPD 120
DB 61 APGEKGEGRGRLPGRCRDPGRGEGAPGPTGPAECVSPPRSASAKRSSESVPPSPD 120
QY 121 APLEPDRVLVNEQCHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESIASFFQ 180
DB 121 APLEPDRVLVNEQCHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESIASFFQ 180
QY 181 FFGGMPKRPASLSGAMVRLPEPDQVWVGVGDYIGITASIKTDSFGFLVYSDMHSSP 240
DB 181 FFGGMPKRPASLSGAMVRLPEPDQVWVGVGDYIGITASIKTDSFGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 9

ID AAB65815 standard; Protein; 243 AA.

AC AAB65815;

DT 28-MAR-2001 (first entry)

XX Human TANGO 253 SEQ ID NO: 3.

XX

KW Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;

KW INTERCEPT 258; coronary disorder; olfactory disorder;

KW neurological disorder; pulmonary disorder; immunological disorder;

KW developmental disorder; kidney disorder.

XX Homo sapiens.

XX WO200078808-A1.

XX 28-DEC-2000.

XX 19-JUN-2000; 2000MO-US16883.

XX 18-JUN-1999; 99US-0336536.

XX (MILL-) MILLENNIUM PHARM INC.

XX Leiby KR, McKay C, Bossone S;

XX WPI: 2001-050109/06.

XX New nucleic acids for treating diseases and disorders, e.g.

XX atherosclerosis, infection, autoimmune diseases, obesity, ear

XX disorders, brain disorders, tumors, diabetes, arthritis, multiple

XX sclerosis and asthma.

XX Claim 9; Page 211-212; 332pp; English.

XX The present invention provides the protein and coding sequences of the
XX human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
XX TANGO 281 and INTERCEPT 258. These are useful in the treatment of
XX coronary, pulmonary, olfactory, immunological, neurological,
XX developmental and kidney disorders.

XX Sequence 243 AA:

Query Match 100.0%; Score 243; DB 22; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.9e-212;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTPGHNGSGLRGDRGDRGAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTPGHNGSGLRGDRGDRGAPG 60
QY 61 APGEKGEGRGRLPGRCRDPGRGEGAPGPTGPAECVSPPRSASAKRSSESVPPSPD 120
DB 61 APGEKGEGRGRLPGRCRDPGRGEGAPGPTGPAECVSPPRSASAKRSSESVPPSPD 120
QY 121 APLEPDRVLVNEQCHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESIASFFQ 180
DB 121 APLEPDRVLVNEQCHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESIASFFQ 180
QY 181 FFGGMPKRPASLSGAMVRLPEPDQVWVGVGDYIGITASIKTDSFGFLVYSDMHSSP 240
DB 181 FFGGMPKRPASLSGAMVRLPEPDQVWVGVGDYIGITASIKTDSFGFLVYSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 10

ID AAB49593 standard; Protein; 243 AA.

AC AAB49593;

DT 13-MAR-2001 (first entry)

XX Human adipocyte complement related protein homolog zs1939.

KW Human; zacrpf5; gene therapy; complement inhibition; C1q domain;

KM adipocyte complement related protein homolog;
 KM inflammation; hormone secretion; inositol phosphate; arachidonate;
 KM phospholipase C activation; gastric emptying; neutrophil activation;
 KM superoxide anion production; antimicrobial; acute vascular injury;
 KM wound healing; zsig39.
 OS
 XX Homo sapiens.
 XX
 PN MO200073444-A1.
 XX
 PD 07-DEC-2000.
 XX
 PF 18-MAY-2000; 2000MO-US13608.
 XX
 PR 27-MAY-1999; 99US-0321372.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Pladdington CS, Sheppard PO;
 DR WPI; 2001-061531/07.
 XX
 PT Novel adipocyte complement related protein homolog, ZACRP5, useful for
 PT diagnosing and treating inflammation, vascular injury microbial
 PT infections, and in wound healing
 XX
 PS Disclosure; Fig 1; 121pp; English.
 XX

CC The present invention relates to human adipocyte complement related
 CC protein homolog, ZACRP5 protein and coding sequence (see AAB49590 and
 CC AAC90045). ZACRP5 has a carboxyl-terminal C1q domain. The ZACRP5 gene is
 CC located on human chromosome 16. ZACRP5 gene and protein are useful for
 CC diagnosing and treating inflammation, for determining arterial
 CC remodeling, for modulating calcium ion concentration, hormone
 CC arachidonate release, phospholipase C activation, gastric emptying, human
 CC neutrophil activation or ADCC capability and superoxide anion production.
 CC ZACRP5 gene and protein are also useful as antimicrobial applications,
 CC preferably against bacteria and virus, for complement inhibition, human
 CC treating acute vascular injury, and for wound healing. The present
 CC sequence is human adipocyte complement related protein homolog, zsig39
 CC protein. This protein was used in a sequence homology alignment with
 CC ZACRP5 protein.
 CC
 XX
 XX
 SQ Sequence 243 AA;

Query Match 100.0%; Score 243; DB 22; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4.9e-212;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCGRHGLGTPGHHGSGQLPGRDGRDGRDGRG 60
 DB 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCGRHGLGTPGHHGSGQLPGRDGRDGRDGRG 60
 QY 61 ARGEKGGGRPGPLGPRGDPGRGEGAPGTPGAGGECVPPRSASFSAKRSESRVPPSD 120
 DB 61 ARGEKGGGRPGPLGPRGDPGRGEGAPGTPGAGGECVPPRSASFSAKRSESRVPPSD 120
 QY 121 APLEFDRLVLENGGHDAVTGKFTCOVPGYVFAVHATVYRASLQFDLVNKGESIASFFQ 180
 DB 121 APLEFDRLVLENGGHDAVTGKFTCOVPGYVFAVHATVYRASLQFDLVNKGESIASFFQ 180
 QY 181 FFGGMPKPAASLSGAMVRLEPEDQVMVGVDYIGIYASIKTDTSTSGFLVYSMDHSSP 240
 DB 181 FFGGMPKPAASLSGAMVRLEPEDQVMVGVDYIGIYASIKTDTSTSGFLVYSMDHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243

RESULT 11
 AAB49599

ID AAB49599 standard; Protein; 243 AA.
 XX
 AC AAB49599;
 XX
 DT 13-MAR-2001 (first entry)
 DE
 XX Human adipocyte complement related protein homolog zsig39.
 XX

KM Human; ZACRP6; gene therapy; complement inhibition; C1q domain;
 KM adipocyte complement related protein homolog;
 KM inflammation; hormone secretion; inositol phosphate; arachidonate;
 KM phospholipase C activation; gastric emptying; neutrophil activation;
 KM superoxide anion production; antimicrobial; acute vascular injury;
 KM wound healing; zsig39.
 OS
 XX Homo sapiens.
 XX
 PN MO200073446-A2.
 XX
 PD 07-DEC-2000.
 XX
 PF 22-MAY-2000; 2000MO-US14024.
 XX
 PR 27-MAY-1999; 99US-0321262.
 XX
 PA (ZYMO) ZYMOGENETICS INC.
 XX
 PI Pladdington CS, Sheppard PO;
 DR WPI; 2001-061532/07.
 XX
 PT Novel adipocyte complement related protein homolog, ZACRP6, useful as
 PT modulators of neurotransmission and for treating disseminated
 PT intravascular coagulation, arteriosclerosis and acute vascular injury
 XX
 PS Disclosure; Fig 1; 119pp; English.
 XX

CC The present invention relates to human adipocyte complement related
 CC protein homolog, ZACRP6 protein and coding sequence (see AAB49596 and
 CC AAC90051). ZACRP6 has a carboxyl-terminal C1q domain. The ZACRP6 gene is
 CC located on human chromosome 21q. ZACRP6 gene and protein are useful for
 CC diagnosing and treating inflammation, for determining arterial
 CC remodeling, for modulating calcium ion concentration, hormone
 CC arachidonate release, phospholipase C activation, gastric emptying, human
 CC neutrophil activation or ADCC capability and superoxide anion production.
 CC ZACRP6 gene and protein are also useful as antimicrobial applications,
 CC preferably against bacteria and virus, for complement inhibition, human
 CC treating acute vascular injury, disseminated intravascular coagulation,
 CC arteriosclerosis and for wound healing. The present sequence is human
 CC adipocyte complement related protein homolog zsig39. This protein was
 CC used in a sequence homology comparison with ZACRP6 protein.
 CC
 XX
 XX
 SQ Sequence 243 AA;

Query Match 100.0%; Score 243; DB 22; Length 243;
 Best Local Similarity 100.0%; Pred. No. 4.9e-212;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCGRHGLGTPGHHGSGQLPGRDGRDGRDGRG 60
 DB 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCGRHGLGTPGHHGSGQLPGRDGRDGRDGRG 60
 QY 61 ARGEKGGGRPGPLGPRGDPGRGEGAPGTPGAGGECVPPRSASFSAKRSESRVPPSD 120
 DB 61 ARGEKGGGRPGPLGPRGDPGRGEGAPGTPGAGGECVPPRSASFSAKRSESRVPPSD 120
 QY 121 APLEFDRLVLENGGHDAVTGKFTCOVPGYVFAVHATVYRASLQFDLVNKGESIASFFQ 180
 DB 121 APLEFDRLVLENGGHDAVTGKFTCOVPGYVFAVHATVYRASLQFDLVNKGESIASFFQ 180
 QY 181 FFGGMPKPAASLSGAMVRLEPEDQVMVGVDYIGIYASIKTDTSTSGFLVYSMDHSSP 240
 DB 181 FFGGMPKPAASLSGAMVRLEPEDQVMVGVDYIGIYASIKTDTSTSGFLVYSMDHSSP 240

Db 61 EAGPAPGTPAGCSVPPRSASFSAKRSERVPSPDAPLDFDRVLVNBQGHDAVTGKFT 120
 QY 145 CQVPGYVFAVHATYRASLOFLVKNESIASFFQFGMPKPSLSGAMVRLPEPDQ 204
 Db 121 CQVPGYVFAVHATYRASLOFLVKNESIASFFQFGMPKPSLSGAMVRLPEPDQ 204
 QY 205 VAVQVGVGDYIGIYASIKTDSFSGFLVYSDMH 237
 Db 181 VAVQVGVGDYIGIYASIKTDSFSGFLVYSDMH 213

RESULT 14

AAB65888
 ID AAB65888 standard; Protein: 243 AA.

AC AAB65888;

DT 28-MAR-2001 (first entry)

DE Human secreted protein related protein SEQ ID NO: 102.

KW Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
 KW INTERCEPT 258; coronary disorder; olfactory disorder;
 KW neurological disorder; pulmonary disorder; immunological disorder;
 KW developmental disorder; kidney disorder.

OS Homo sapiens.

PN MO200078808-A1.

PD 28-DEC-2000.

PE 19-JUN-2000; 2000MO-US16883.

PR 18-JUN-1999; 99US-0336536.

PA (MILL-) MILLENNIUM PHARM INC.

PI Leiby KR, McKay C, Bossone S;

DR WPI: 2001-050109/06.

XX New nucleic acids for treating diseases and disorders, e.g.

XX atherosclerosis, infection, autoimmune diseases, obesity, ear

XX disorders, brain disorders, tumors, diabetes, arthritis, multiple

XX sclerosis and asthma

XX Disclosure: Page 270-271; 332pp; English.

XX The present invention provides the protein and coding sequences of the

XX human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,

XX TANGO 281 and INTERCEPT 258. These are useful in the treatment of

XX coronary, pulmonary, olfactory, immunological, neurological,

XX developmental and kidney disorders.

XX Sequence 243 AA;

XX Query Match

XX Best Local Similarity 76.18; Score 185; DB 22; Length 243;

XX Matches 185; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 PGAPGEGGRCGLPGRCGDPGRGEGAGPAGTGPAGCSVPPRSASFSAKRSERVP 118
 Db 59 PGAPGEGGRCGLPGRCGDPGRGEGAGPAGTGPAGCSVPPRSASFSAKRSERVP 118
 QY 119 SDAPLDFDVLVNBQGHDAVTGKFTCOVPGYVFAVHATYRASLOFLVKNESIASF 178
 Db 119 SDAPLDFDVLVNBQGHDAVTGKFTCOVPGYVFAVHATYRASLOFLVKNESIASF 178
 QY 179 FQFGGMPKPSLSGAMVRLPEPDQVAVQVGVGDYIGIYASIKTDSFSGFLVYSDMH 238
 Db 179 FQFGGMPKPSLSGAMVRLPEPDQVAVQVGVGDYIGIYASIKTDSFSGFLVYSDMH 238

QY 239 SPVFA 243
 Db 239 SPVFA 243

RESULT 15

AAB65891

ID AAB65891 standard; Protein: 243 AA.

AC AAB65891;

DT 28-MAR-2001 (first entry)

DE Human secreted protein related protein SEQ ID NO: 108.

KW Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
 KW INTERCEPT 258; coronary disorder; olfactory disorder;
 KW neurological disorder; pulmonary disorder; immunological disorder;
 KW developmental disorder; kidney disorder.

OS Homo sapiens.

PN MO200078808-A1.

PD 28-DEC-2000.

PE 19-JUN-2000; 2000MO-US16883.

PR 18-JUN-1999; 99US-0336536.

PA (MILL-) MILLENNIUM PHARM INC.

PI Leiby KR, McKay C, Bossone S;

DR WPI: 2001-050109/06.

XX New nucleic acids for treating diseases and disorders, e.g.

XX atherosclerosis, infection, autoimmune diseases, obesity, ear

XX disorders, brain disorders, tumors, diabetes, arthritis, multiple

XX sclerosis and asthma

XX Disclosure: Page 274; 332pp; English.

XX The present invention provides the protein and coding sequences of the

XX human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,

XX TANGO 281 and INTERCEPT 258. These are useful in the treatment of

XX coronary, pulmonary, olfactory, immunological, neurological,

XX developmental and kidney disorders.

XX Sequence 243 AA;

XX Query Match

XX Best Local Similarity 71.68; Score 174; DB 22; Length 243;

XX Matches 174; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCGHPGLPGRPHHSGSLPRDRGDRGAG 60
 Db 1 MRPLVLLLLGLAAGSPPLDNNKIPSLCGHPGLPGRPHHSGSLPRDRGDRGAG 60
 QY 61 APGEKGGRCGLPGRCGDPGRGEGAGPAGTGPAGCSVPPRSASFSAKRSERVP 120
 Db 61 APGEKGGRCGLPGRCGDPGRGEGAGPAGTGPAGCSVPPRSASFSAKRSERVP 120
 QY 121 APPLPFDRLVNBQGHDAVTGKFTCOVPGYVFAVHATYRASLOFLVKNES 174
 Db 121 APPLPFDRLVNBQGHDAVTGKFTCOVPGYVFAVHATYRASLOFLVKNES 174

Search completed: June 18, 2003, 15:03:51
 Job time : 77 secs

status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA
 A:Residues: 1-289 <MIL>
 A:Cross-references: EMBL:U55373; PIDN:AAC25888.1; GSPDB:GN00023; CESP:F26F12.1
 A:Experimental source: stratin Bristol N2; clone F26F12
 C:Genetics:
 A:Gene: CESP:F26F12.1
 A:Map position: 5
 A:Introns: 45/3
 C:Superfamily: unassigned collagens

Query Match
 Best Local Similarity 100.0%; Score 11; DB 2; Length 289;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 80 PGPGEAGPAG 90
 DB 222 PGPGEAGPAG 232

RESULT 3

Surfactant protein D - bovine
 C:Species: Bos primigenius taurus (cattle)
 C>Date: 02-Dec-1993 #sequence-revision 01-Sep-1995 #text-change 17-Mar-1999
 C:Accession: S33603
 R:Lim, B.L.; Lu, J.; Reid, K.B.M.
 Immunology 78, 159-165, 1993
 A:Title: Structural similarity between bovine conglutinin and bovine lung surfactant pro
 A:Reference number: S33603; MUID:93170856; PMID:8436402
 A:Accession: S33603
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-369 <LIM>
 C:Superfamily: pulmonary surfactant protein D; C-type lectin homology
 F:248-367/Domain: C-type lectin homology <LCH>

Query Match
 Best Local Similarity 100.0%; Score 11; DB 2; Length 369;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRGDRGR 55
 DB 46 GLPGRGDRGR 56

RESULT 4

Surfactant protein D - rat
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 10-Sep-1999 #sequence-revision 10-Sep-1999 #text-change 10-Sep-1999
 C:Accession: A42046
 R:Shimizu, H.; Fisher, J.H.; Papst, P.; Benson, B.; Lau, K.; Mason, R.J.; Voelker, D.R.
 J. Biol. Chem. 267, 1853-1857, 1992
 A:Title: Primary structure of rat pulmonary surfactant protein D. cDNA and deduced amino
 A:Reference number: A42046; MUID:93112913; PMID:1370483
 A:Accession: A42046
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-374 <SHI>
 A:Cross-references: PIDN:AAB25037.1; PID:9263973; PIDN:AAB25038.1; PID:9263974
 A:Experimental source: lung
 A:Note: sequence extracted from NCBI backbone (NCBI:76027, NCBI:76031)
 C:Superfamily: pulmonary surfactant protein D; C-type lectin homology
 F:253-372/Domain: C-type lectin homology <LCH>

Query Match
 Best Local Similarity 100.0%; Score 11; DB 1; Length 374;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRGDRGR 55
 DB 45 GLPGRGDRGR 55

RESULT 5

Surfactant protein D precursor - human
 A:Accession: A45225
 N:Alternate names: SP-D
 C:Species: Homo sapiens (man)
 C>Date: 16-Apr-1999 #sequence-revision 16-Apr-1999 #text-change 22-Jun-1999
 C:Accession: A45225; S23434; S24555; S44420; S18382; A56776
 R:Crouch, E.; Rust, K.; Velle, R.; Donis-Keller, H.; Grosso, L.
 J. Biol. Chem. 268, 2976-2983, 1993
 A:Title: Genomic organization of human surfactant protein D (SP-D). SP-D is encoded o
 A:Reference number: A45225; MUID:93155122; PMID:8428971
 A:Accession: A45225
 A:Molecule type: DNA
 A:Residues: 1-375 <CRO>
 A:Cross-references: GB:L05483; GB:L05484; GB:L05485; NID:9292505; PIDN:AAB59450.1; PI
 A:Experimental source: placenta
 A:Note: sequence extracted from NCBI backbone (NCBI:124316)
 R:Lu, J.; Willis, A.C.; Reid, K.B.M.
 Biochem. J. 284, 795-802, 1992
 A:Title: Purification, characterization and cDNA cloning of human lung surfactant pro
 A:Reference number: S23434; MUID:92322003; PMID:1339284
 A:Accession: S23434
 A:Molecule type: mRNA
 A:Residues: 1-30, 'T', '32-121, 'P', '123-179, 'A', '181-375 <LUJ1>
 A:Cross-references: EMBL:X65018; NID:934766; PIDN:CAA46152.1; PID:934767
 A:Experimental source: lung
 A:Accession: S24555
 A:Molecule type: protein
 A:Residues: 214-234, 'X', '236, 'XX', '239-241 <LUJ2>
 R:Hoppe, H.J.; Barlow, P.N.; Reid, K.B.M.
 FEBS Lett. 344, 191-195, 1994
 A:Title: A parallel three stranded alpha-helical bundle at the nucleation site of col
 A:Reference number: S44420; MUID:94244769; PMID:8187882
 A:Accession: S44420
 A:Molecule type: mRNA
 A:Residues: 202-257 <HOP>
 R:Rust, K.; Grosso, L.; Zhang, V.; Chang, D.; Persson, A.; Longmore, W.; Cal, G.Z.; C
 Arch. Biochem. Biophys. 290, 116-126, 1991
 A:Title: Human surfactant protein D: SP-D contains a C-type lectin carbohydrate recog
 A:Reference number: S18382; MUID:91378578; PMID:1898081
 A:Accession: S18382
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 'F', '60-205, 'P', '207-374, 'HF', <RUS>
 A:Cross-references: GB:L05485; NID:9292505
 A:Note: corrections to this sequence are reported in reference A56776
 R:Crouch, E.; Persson, A.; Chang, D.
 Am. J. Pathol. 142, 241-248, 1993
 A:Title: Accumulation of surfactant protein D in human pulmonary alveolar proteinosis
 A:Reference number: A56776; MUID:93142849; PMID:8424457
 A:Accession: A56776
 A:Status: preliminary
 A:Molecule type: protein
 A:Residues: 46-58, 'F', '60-62, 'E', '64-72, '223-227, 'X', '229-239, 'P', '241-245, 'X', '247-256, 'X'
 A:Cross-references: PIDN:AAB25037.1; PID:9263973; PIDN:AAB25038.1; PID:9263974
 A:Experimental source: bronchoalveolar lavage
 A:Note: sequence extracted from NCBI backbone (NCBI:123024, NCBI:123023); sequence
 C:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lower
 A:Gene: GDB:SFPP4; SP-D
 A:Cross-references: GDB:132674; OMIM:178635
 A:Map position: 10q22.2-10q23.1
 C:Superfamily: pulmonary surfactant protein D; C-type lectin homology
 C:Keywords: blocked amino end; calcium; glycoprotein; hydroxylsine; hydroxyproline;
 F:1-10/Domain: signal sequence #status predicted <SIG>
 F:21-375/Product: pulmonary surfactant protein D #status predicted <MAT>
 F:21-45/Domain: non-collagenous #status predicted <NC1>
 F:46-222/Domain: non-collagenous #status predicted <MAT>
 F:223-375/Domain: non-collagenous #status predicted <COL>
 F:254-373/Domain: C-type lectin homology <LCH>
 F:90/Binding site: carbohydrate (asn) (covalent) #status predicted

F:281-373,351-365/Disulfide bonds: #status predicted

Query Match 4.5%; Score 11; DB 1; Length 375;
Best Local Similarity 100.0%; Pred. No. 0.021;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDGR 55
|||||
Db 46 GLPGRDGRDGR 56

RESULT 6

CGHUB3
collagen alpha 3(IV) chain precursor, long splice form - human

M:Alternate names: Goodpasture antigen; procollagen alpha 3(IV) chain long splice form
C:Species: Homo sapiens (man)
C:Date: 28-Oct-1994 #sequence, revision 03-Oct-1995 #text_change 22-Jun-1999

C:Accession: A54763; #sequence, revision 03-Oct-1995 #text_change 22-Jun-1999
R:Matsumoto, M.; Leinonen, A.; Mochizuki, T.; Tryggvason, K.; Reiders, S.T.
J. Biol. Chem. 269, 23013-23017, 1994

A:Title: Complete primary structure of the human alpha3(IV) collagen chain. Coexpressor
A:Reference number: A54763; MUID:94364994; PMID:8083201

A:Accession: A54763
A:Molecule type: mRNA

A:Residues: 1-1670 <MAR>

A:Cross-references: GB:X80031; NID:9577563; PID:9577564
A:Experimental source: Kidney

R:Turner, N.; Mason, P.J.; Brown, R.; Fox, M.; Povey, S.; Rees, A.; Pusey, C.D.
J. Clin. Invest. 89, 592-601, 1992

A:Title: Molecular cloning of the human Goodpasture antigen demonstrates it to be the alpha
A:Reference number: A43928; MUID:92147878; PMID:1737849

A:Accession: A43928
A:Molecule type: mRNA

A:Residues: 1331-1524, '1', 1526-1670 <TUR>

A:Cross-references: GB:M1379
A:Experimental source: Kidney

R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J. Biol. Chem. 267, 19780-19784, 1992

A:Title: Exon/intron structure of the human alpha 3(IV) gene encompassing the Goodpasture
A:Reference number: A44043; MUID:93015826; PMID:1400291

A:Accession: A44043
A:Molecule type: DNA; mRNA

A:Residues: 1386-1670 <OUT>

A:Cross-references: GB:M92993; NID:9177895; PID:AAA21610.1; PID:9177896
A:Note: Sequence extracted from NCBI backbone (NCBI:115597)

R:Quinones, S.; Bernal, D.; Garcia-Sogo, M.; Elena, S.F.; Saus, J.
J. Biol. Chem. 269, 17358, 1994

A:Reference number: A44738; MUID:94274734; PMID:8006044
A:Contents: annotation; erratum; correction to Intronic sequence in A44043

R:Bernal, D.; Quinones, S.; Saus, J.
J. Biol. Chem. 268, 12090-12094, 1993

A:Title: The human mRNA encoding the Goodpasture antigen is alternatively spliced.
A:Reference number: A45971; MUID:93280184; PMID:8505332

A:Accession: A45971
A:Molecule type: mRNA

A:Status: nucleic acid sequence not shown
A:Note: sequence extracted from NCBI backbone (NCBI:133363); sequence incorrectly identified

R:Morrison, K.E.; Matsumoto, M.; Yang-Feng, T.L.; Reiders, S.T.
Am. J. Hum. Genet. 49, 545-554, 1991

A:Title: Sequence and localization of a partial cDNA encoding the human alpha3 chain of
A:Reference number: A39786; MUID:91535370; PMID:1882840

A:Accession: A39786
A:Molecule type: mRNA

A:Residues: 1453-1593, 'A', 1595-1670 <MOR>

A:Cross-references: GB:S55790; NID:9234418; PID:AA19637.1; PID:9234419
C:Comment: Prolines and lysines at the third position of the tripeptide repeating unit
C:Comment: In Goodpasture's syndrome, an autoimmune response develops against an epitope

A:Gene: GDB:COL4A3
A:Cross-references: GDB:128351; OMIM:120070

A:Map position: 2q36-2q37

A:Introns: 1385/1; 1418/1; 1488/1; 1547/2; 1585/3; 1643/2 #status incomplete

A:Note: the alpha 3(IV) and alpha 4(IV) chain genes are encoded on opposite strands w
C:Complex: This minor type IV collagen is thought to form a heterotrimer of two alpha
monomer amino-terminal domains (with disulfide and desmosine cross-links), dimer
er associations in the interrupted helical domain (with disulfide and desmosine cross

C:Function:
A:Description: minor structural component of extracellular basement membrane in kidney

C:Superfamily: collagen alpha 1(IV) chain

C:Keywords: alternative splicing; basement membrane; cell binding; coiled coil; extra

F:1-28/Domain: signal sequence #status predicted <SIG>

F:29-42/Domain: amino-terminal nonhelical, NHI <NHI>

F:43-1438/Region: interrupted helical

F:791-793/Region: cell attachment (R-G-D) motif

F:996-998/Region: cell attachment (R-G-D) motif

F:1154-1156/Region: cell attachment (R-G-D) motif

F:1306-1308/Region: cell attachment (R-G-D) motif

F:1345-1347/Region: cell attachment (R-G-D) motif

F:1432-1434/Region: cell attachment (R-G-D) motif

F:1439-1670/Domain: carboxyl-terminal nonhelical, NCI <NCI>

F:1451-1551/Domain: collagen IV carboxyl-terminal repeat <CT1>

F:31,33,39,41,125,422,476,479,682,722,809,1387/Disulfide bonds: Interchain #status pr

F:253/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:1460-1548,1493-1551/Disulfide bonds: (or 1460-1551, 1493-1548) #status predicted

F:1505-1511,1616-1622/Disulfide bonds: #status predicted

F:1570-1662,1604-1665/Disulfide bonds: (or 1570-1665, 1604-1662) #status predicted

Query Match

Best Local Similarity 4.5%; Score 11; DB 1; Length 1670;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 71 PGLPGRDGRDGR 81
|||||
Db 1300 PGLPGRDGRDGR 1310

RESULT 7

A54849
collagen alpha 1(VII) chain precursor - human

M:Alternate names: procollagen alpha 1(VII) chain
C:Species: Homo sapiens (man)

C:Date: 04-Nov-1994 #sequence, revision 04-Nov-1994 #text_change 20-Sep-1999
C:Accession: A54849; PH0844; S16316; 156328; A30296; 184666

R:Christiano, A.M.; Greenspan, D.S.; Lee, S.; Uitto, J.
J. Biol. Chem. 269, 20256-20262, 1994

A:Title: Cloning of human type VII collagen. Complete primary sequence of the alpha1
A:Reference number: A54849; MUID:94327588; PMID:8051117

A:Accession: A54849
A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-2944 <CHR>

A:Cross-references: GB:L02870; NID:9987124; PID:AAA75438.1; PID:9987125
R:Tanaka, T.; Takahashi, K.; Furukawa, F.; Imanura, S.
Biochem. Biophys. Res. Commun. 183, 958-963, 1992

A:Title: Molecular cloning and characterization of type VII collagen cDNA.
A:Reference number: PH0844; MUID:92231902; PMID:1567409

A:Accession: PH0844
A:Molecule type: mRNA

A:Residues: 'EPR', 340-475, 'RALSTASHSTLCMRATRMHPCNRGSHWTRAAECGCPNPASRAARAG', 524-528,
A:Cross-references: DDBJ:D11152; DDBJ:D13694; NID:9453698; PID:BA02853.1; PID:94536

A:Experimental source: keratinocyte

A:Note: the authors translated the codon ACC for residues 394 and 397 as Tyr
Proc. Natl. Acad. Sci. U.S.A. 88, 6931-6935, 1991

A:Title: Human type VII collagen: cDNA cloning and chromosomal mapping of the gene.
A:Reference number: S16316; MUID:91334380; PMID:1871109

A:Accession: S16316
A:Molecule type: mRNA

A:Residues: 815-892, 'E', 894-1439 <PAR>

A:Cross-references: GB:M65158; GB:S45017; NID:9180914; PID:AAA96439.1; PID:9180915
A:Experimental source: keratinocyte

R:Gannon, W.R.; Abernethy, M.L.; Padilla, K.M.; Prisaanah, P.S.; Cook, M.E.; Wright, J.
 J. Invest. Dermatol. 99, 691-696, 1992
 A:Title: Noncollagenous (NC1) domain of collagen VII resembles multidomain adhesion prot
 A:Reference number: 156328; MUID:93107742; PMID:1469284
 A:Accession: 156328
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 'EPR', 372-517, 'DV', 520-540, 'W', 542-1255 <RES>
 A:Cross-references: GB:551235; NID:9262308; PIDN:AA89196.1; PID:9262309
 R:Selzter, J.L.; Eisen, A.Z.; Bauer, E.A.; Morris, N.P.; Glanville, R.W.; Burgeson, R.E.
 J. Biol. Chem. 264, 3822-3826, 1989
 A:Title: Cleavage of type VII collagen by interstitial collagenase and type IV collagen
 A:Reference number: A30296; MUID:89139437; PMID:2537292
 A:Accession: A30296
 A:Molecule type: protein
 A:Residues: A, 1240-1246, 'G', 1248-1250, 'XE', 1253-1255, 'Q', 1257, 'E', 2032, 'C', 2034-2041,
 A:Note: Two reported peptides cannot be reliably located
 R:Greene, D.S.
 Hum. Mol. Genet. 2, 273-278, 1993
 A:Title: The carboxyl-terminal half of type VII collagen, including the non-collagenous
 A:Reference number: 148103; MUID:93271985; PMID:8499916
 A:Accession: 148103
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 2395-2871, 'S', 2873-2944 <RE2>
 A:Cross-references: GB:106862; NID:9388713; PIDN:AA89196.1; PID:9388714
 R:Christiano, A.M.; Rymaszewski, M.; Uitto, J.
 Proc. Natl. Acad. Sci. U.S.A. 91, 3549-3553, 1994
 A:Title: Dominant dystrophic epidermolysis bullosa: identification of a Gly --> Ser subs
 A:Reference number: A55255; MUID:94224777; PMID:8170945
 A:Accession: A55255
 A:Contents: annotation
 C:Comment: prolines and lysines at the third position of the tripeptide repeating unit
 C:Comment: and subsequently O-glycosylated.
 C:Genetics:
 A:Gene: GDB:COL7A1; EBR1; EBD1; EB
 A:Cross-references: GDB:128750; OMIM:120120
 A:Map position: 3p21.3-3p21.3
 A:Note: defects in this gene can result in dominant and recessive dystrophic epidermolys
 A:Note: there are 118 introns
 C:Complex: type VII collagen is probably a homotrimer
 C:Function:
 A:Description: structural component of extracellular polymer associated with anchoring
 C:Superfamily: unassigned collagens; animal knittz-type proteinase inhibitor
 C:Keywords: coll; extracellular matrix; glycoprotein; hydroxyllysine; hydroxyproli
 F:1-16/Domain: signal sequence #status predicted <SIG>
 F:17-2944/Product: collagen alpha 1(VII) chain #status predicted <MAT>
 F:17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>
 F:17-1253/Domain: amino-terminal nonhelical #status predicted <NC1>
 F:36-201/Domain: von Willibrand factor type A repeat homology <VMA1>
 F:327-413/Domain: fibronectin type III repeat homology <FN1>
 F:414-502/Domain: fibronectin type III repeat homology <FN2>
 F:508-593/Domain: fibronectin type III repeat homology <FN3>
 F:598-683/Domain: fibronectin type III repeat homology <FN4>
 F:686-771/Domain: fibronectin type III repeat homology <FN5>
 F:776-862/Domain: fibronectin type III repeat homology <FN6>
 F:864-952/Domain: fibronectin type III repeat homology <FN7>
 F:954-1045/Domain: fibronectin type III repeat homology <FN8>
 F:1052-1119/Domain: von Willibrand factor type A repeat homology <VMA2>
 F:1170-1172/Domain: von Willibrand factor type A repeat homology <VMA2>
 F:1189-1253/Region: cysteine/proline-rich
 F:1254-2783/Region: interrupted helical
 F:1334-1336/Region: cell attachment (R-G-D) motif
 F:2008-2010/Region: cell attachment (R-G-D) motif
 F:2553-2555/Region: cell attachment (R-G-D) motif
 F:2784-2944/Domain: carboxyl-terminal nonhelical #status predicted <NC2>
 F:2876-2929/Domain: animal knittz-type proteinase inhibitor homology <BTR>
 F:337-786, 1109/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:2167, 2176, 2188, 2664, 2667, 2673/Modified site: 4-hydroxyproline (Pro) #status exp
 F:2623, 2631/Modified site: 5-hydroxylysine (Lys) (covalent) #status experimental
 F:2633, 2631/Binding site: carbohydrate (Lys) (covalent) #status experimental
 F:2634, 2802, 2804/Disulfide Bonds: Interchain #status predicted

Query Match 4.5% Score 11; DB 2; Length 2944;

Best Local Similarity 100.0%; Pred. No. 0.12;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 71 GGLPGRGDPG 81
 DB 1328 GGLPGRGDPG 1338

RESULT 8
 B44984
 collagen - nematode (Haemonchus contortus) (fragment)
 C:Species: Haemonchus contortus
 C:Date: 28-Apr-1993 #sequence_revision 28-Apr-1993 #text_change 07-May-1999
 C:Accession: B44984
 R:Shamsky, L.M.; Pratt, D.; Bolsvenne, R.J.; Cox, G.N.
 Mol. Biochem. Parasitol. 37, 73-86, 1989
 A:Title: Cuticle collagen genes of Haemonchus contortus and Caenorhabditis elegans ar
 A:Reference number: A44984; MUID:90136718; PMID:2615789
 A:Accession: B44984
 A:Status: preliminary
 A:Molecule type: mRNA
 A:Residues: 1-210 <SHA>
 A:Cross-references: GB:J04671; GB:J04670
 C:Superfamily: unassigned collagens

Query Match 4.1% Score 10; DB 2; Length 210;
 Best Local Similarity 100.0%; Pred. No. 0.11;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 57 GARGPGERG 66
 DB 151 GARGPGERG 160

RESULT 9
 LNRBS
 pulmonary surfactant protein A precursor - rabbit
 N:Alternate names: pulmonary surfactant 32k apoprotein; pulmonary surfactant-associat
 C:Species: Oryctolagus cuniculus (domestic rabbit)
 C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999
 C:Accession: A29931
 R:Bogart, V.; Qing, K.; Mendelson, C.R.
 J. Biol. Chem. 263, 2939-2947, 1988
 A:Title: The major apoprotein of rabbit pulmonary surfactant. Elucidation of primary
 A:Reference number: A29931; MUID:88139348; PMID:2830270
 A:Accession: A29931
 A:Molecule type: mRNA
 A:Residues: 1-247 <BOG>
 A:Cross-references: GB:J03542; NID:9165705; PIDN:AA31465.1; PID:9165706
 A:Note: 12-Pro was also found
 A:Note: two species of mRNA, which appear to be transcribed from a single gene, could
 A:Note: the amino end of the mature protein is blocked
 C:Comment: pulmonary surfactant is a complex of phospholipids and proteins that lower
 C:Comment: This protein is a sialoglycoprotein synthesized by alveolar type II cells.
 C:Superfamily: mannose-binding lectin; C-type lectin homology
 C:Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; gl
 F:1-15/Domain: signal sequence #status predicted <SIG>
 F:16-247/Product: pulmonary surfactant protein A #status predicted <MAT>
 F:27-99/Region: collagen-like
 F:126-245/Domain: C-type lectin homology <LCH>
 F:126-245/Domain: C-type lectin homology <LCH> (in mature form) #status predicted
 F:16/Modified site: acetylated amino end (Ser) (covalent) #status predicted
 F:206/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.1% Score 10; DB 1; Length 247;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 45 GGLPGRGDPG 54
 DB 39 GGLPGRGDPG 48

RESULT 10

LNHUP6

N:Alternate names: pulmonary surfactant precursor (clone 6A) - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999

C:Accession: A25720

R:Floros, J.; Steinhilber, R.; Jacobs, K.; Phelps, D.; Kriz, R.; Recny, M.; Sultzman, L.;

J. Biol. Chem. 261, 9029-9033, 1986

A:Title: Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfactant

A:Reference number: A25720; MUID:86250832; PMID:3755136

A:Accession: A25720

A:Molecule type: mRNA

A:Residues: 1-248 <PRO>

A:Cross-references: GB:M13686; NID:9190669; PID:AA60211.1; PID:9190670

A:Note: part of the sequence was confirmed by protein sequencing

A:Note: the amino end of the mature protein, which was not identified, is partially acet

A:Note: clones corresponding to two different proteins were sequenced. Cotranslational m

C:GeneID: 825720

A:Gene: GDB:SFRP1; SFRP1; SP-A; SP-A1

A:Cross-references: GDB:119593; OMIM:178630

A:Map position: 10q22-10q23

C:Superfamily: mannose-binding lectin; C-type lectin homology

C:Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; glyco

F:1-20/Domain: signal sequence #status predicted <SIG>

F:127-246/Product: pulmonary surfactant protein A #status predicted <MAT>

F:21/Modified site: acetylated amino end (Glu) (in mature form) #status predicted

F:30,33,36,42,54,57,63,76,79,82,91,97/Modified site: 4-hydroxyproline (Pro) #status pred

F:207/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 4.1%; Score 10; DB 1; Length 248;

Best Local Similarity 100.0%; Pred. No. 0.13;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDG 54

DB 40 GLPGRDGRDG 49

LNHUP1

N:Alternate names: pulmonary surfactant precursor (clone 1A) - human

C:Species: Homo sapiens (man)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999

C:Accession: B25720

R:Floros, J.; Steinhilber, R.; Jacobs, K.; Phelps, D.; Kriz, R.; Recny, M.; Sultzman, L.;

J. Biol. Chem. 261, 9029-9033, 1986

A:Title: Isolation and characterization of cDNA clones for the 35-kDa pulmonary surfact

A:Reference number: A25720; MUID:86250832; PMID:3755136

A:Accession: B25720

A:Molecule type: mRNA

A:Residues: 1-248 <PRO>

A:Cross-references: GB:K03475

A:Note: part of the sequence was confirmed by protein sequencing

A:Note: the amino end of the mature protein, which was not identified, is partially acet

A:Note: clones corresponding to two different proteins were sequenced. Cotranslational m

C:GeneID: 825720

A:Gene: GDB:SFRP1; SFRP1; SP-A; SP-A1

A:Cross-references: GDB:119593; OMIM:178630

A:Map position: 10q22-10q23

C:Superfamily: mannose-binding lectin; C-type lectin homology

C:Keywords: acetylated amino end; alveolar proteinosis; calcium; gaseous exchange; glyco

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDG 54

DB 40 GLPGRDGRDG 49

LNHUP6

N:Alternate names: pulmonary surfactant precursor (clone 6A) - human

C:Species: Canis lupus familiaris (dog)

C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999

C:Accession: A25296; A61227; A60142

R:Benson, B.; Hawgood, S.; Schilling, J.; Clements, J.; Damm, D.; Cordell, B.; White,

Proc. Natl. Acad. Sci. U.S.A. 82, 6379-6383, 1985

A:Title: Structure of canine pulmonary surfactant apoprotein: cDNA and complete amino

A:Reference number: A25296; MUID:86016705; PMID:3863100

A:Accession: A25296

A:Molecule type: mRNA

A:Residues: 1-248 <BEN>

A:Note: the authors translated the codon TTC for residue 60 as Pro

A:Note: part of the sequence, including the amino end of the mature protein, was conf

R:Liaw, D.F.; Ryan, S.F.

Chem. Phys. Lipids 59, 29-38, 1991

A:Title: Purification of surfactant protein A from dog lung by reconstitution with su

A:Reference number: A61227; MUID:92163993; PMID:11790579

A:Accession: A61227

A:Molecule type: protein

A:Residues: 18-32 <LIA>

R:Ross, G.F.; Meuth, J.; Ohlring, B.; Kim, Y.; Whitsett, J.A.

Biochim. Biophys. Acta 870, 267-278, 1986

A:Title: Purification of canine surfactant-associated glycoproteins A. Identification

A:Reference number: A60142; MUID:86159848; PMID:3006781

A:Accession: A60142

A:Molecule type: protein

A:Residues: 24-34; 95-101, 'X', 103-108 <ROS>

R:Patthy, L.

Nature 325, 490, 1987

A:Reference number: A93388; MUID:87115834; PMID:3808053

A:Contents: annotation: animal lectin domain homology

C:Comment: This protein is a complex of phospholipids and proteins that lower

pendent on the presence of calcium ions.

C:Superfamily: mannose-binding lectin; C-type lectin homology

C:Keywords: alveolar proteinosis; calcium; disulfide bond; gaseous exchange; hydroxyp

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-248/Product: pulmonary surfactant protein A #status experimental <MPT>

F:28-102/Region: collagen-like

F:127-246/Domain: C-type lectin homology <LCH>

F:20,207/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:30/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 4.1%; Score 10; DB 1; Length 248;

Best Local Similarity 100.0%; Pred. No. 0.13;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDG 54

DB 40 GLPGRDGRDG 49

LNHUP6

N:Alternate names: pulmonary surfactant precursor (clone 6A) - human

C:Species: Canis lupus familiaris (dog)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999

C:Accession: A25296; A61227; A60142

R:Benson, B.; Hawgood, S.; Schilling, J.; Clements, J.; Damm, D.; Cordell, B.; White,

Proc. Natl. Acad. Sci. U.S.A. 82, 6379-6383, 1985

A:Title: Structure of canine pulmonary surfactant apoprotein: cDNA and complete amino

A:Reference number: A25296; MUID:87213191; PMID:3579914

A:Accession: A25296

A:Molecule type: mRNA

A:Residues: 1-248 <BEN>

A:Note: the authors translated the codon TTC for residue 60 as Pro

A:Note: part of the sequence, including the amino end of the mature protein, was conf

R:Liaw, D.F.; Ryan, S.F.

Chem. Phys. Lipids 59, 29-38, 1991

A:Title: Purification of surfactant protein A from dog lung by reconstitution with su

A:Reference number: A61227; MUID:92163993; PMID:11790579

A:Accession: A61227

A:Molecule type: protein

A:Residues: 18-32 <LIA>

R:Ross, G.F.; Meuth, J.; Ohlring, B.; Kim, Y.; Whitsett, J.A.

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDG 54

DB 40 GLPGRDGRDG 49

LNHUP6

N:Alternate names: pulmonary surfactant precursor (clone 6A) - human

C:Species: Canis lupus familiaris (dog)

C:Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 16-Jul-1999

C:Accession: A25296; A61227; A60142

R:Benson, B.; Hawgood, S.; Schilling, J.; Clements, J.; Damm, D.; Cordell, B.; White,

Proc. Natl. Acad. Sci. U.S.A. 82, 6379-6383, 1985

A:Title: Structure of canine pulmonary surfactant apoprotein: cDNA and complete amino

A:Reference number: A25296; MUID:86016705; PMID:3863100

A:Accession: A25296

A:Molecule type: mRNA

A:Residues: 1-248 <BEN>

A:Note: the authors translated the codon TTC for residue 60 as Pro

A:Note: part of the sequence, including the amino end of the mature protein, was conf

R:Liaw, D.F.; Ryan, S.F.

Chem. Phys. Lipids 59, 29-38, 1991

A:Title: Purification of surfactant protein A from dog lung by reconstitution with su

A:Reference number: A61227; MUID:92163993; PMID:11790579

A:Accession: A61227

A:Molecule type: protein

A:Residues: 18-32 <LIA>

R:Ross, G.F.; Meuth, J.; Ohlring, B.; Kim, Y.; Whitsett, J.A.

Biochim. Biophys. Acta 870, 267-278, 1986

A:Title: Purification of canine surfactant-associated glycoproteins A. Identification

A:Reference number: A60142; MUID:86159848; PMID:3006781

A:Accession: A60142

A:Molecule type: protein

A:Residues: 24-34; 95-101, 'X', 103-108 <ROS>

R:Patthy, L.

Nature 325, 490, 1987

A:Reference number: A93388; MUID:87115834; PMID:3808053

A:Contents: annotation: animal lectin domain homology

C:Comment: This protein is a complex of phospholipids and proteins that lower

pendent on the presence of calcium ions.

C:Superfamily: mannose-binding lectin; C-type lectin homology

C:Keywords: alveolar proteinosis; calcium; disulfide bond; gaseous exchange; hydroxyp

F:1-17/Domain: signal sequence #status predicted <SIG>

F:18-248/Product: pulmonary surfactant protein A #status experimental <MPT>

F:28-102/Region: collagen-like

F:127-246/Domain: C-type lectin homology <LCH>

F:20,207/Binding site: carbohydrate (Asn) (covalent) #status predicted

F:30/Modified site: 4-hydroxyproline (Pro) #status experimental

Query Match 4.1%; Score 10; DB 1; Length 248;

Best Local Similarity 100.0%; Pred. No. 0.13;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDG 54

DB 40 GLPGRDGRDG 49

LNHUP6

N:Alternate names: pulmonary surfactant precursor (clone 6A) - human

C:Species: Canis lupus familiaris (dog)

C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 16-Jul-1999

C:Accession: A25296; A61227; A60142

R:Benson, B.; Hawgood, S.; Schilling, J.; Clements, J.; Damm, D.; Cordell, B.; White,

Proc. Natl. Acad. Sci. U.S.A. 82, 6379-6383, 1985

A:Title: Structure of canine pulmonary surfactant apoprotein: cDNA and complete amino

A:Reference number: A25296; MUID:87213191; PMID:3579914

A:Accession: A25296

A:Molecule type: mRNA

A:Residues: 1-248 <BEN>

A:Note: the authors translated the codon TTC for residue 60 as Pro

A:Note: part of the sequence, including the amino end of the mature protein, was conf

R:Liaw, D.F.; Ryan, S.F.

Chem. Phys. Lipids 59, 29-38, 1991

A:Title: Purification of surfactant protein A from dog lung by reconstitution with su

A:Reference number: A61227; MUID:92163993; PMID:11790579

A:Accession: A61227

A:Molecule type: protein

A:Residues: 18-32 <LIA>

R:Ross, G.F.; Meuth, J.; Ohlring, B.; Kim, Y.; Whitsett, J.A.

A:Contents: 0.9 kb cDNA
 A:Accession: A29299
 A:Molecule type: mRNA
 A:Residues: 1-248 <SAN>
 A>Note: part of the sequence, including the amino end of the mature protein, was confirmed by Fisher, J.H.; Emile, P.A.; Shannon, J.; Sano, K.; Hattler, B.; Mason, R.J.
 R:Fisher, J.H.; Emile, P.A.; Shannon, J.; Sano, K.; Hattler, B.; Mason, R.J.
 A:Title: Rat pulmonary surfactant protein A is expressed as two differently sized mRNA
 A:Reference number: J50034; MUID:89000785; PMID:2901856
 A:Contents: 1.6 kb cDNA
 A:Accession: J50034
 A:Molecule type: mRNA
 A:Residues: 1-248 <FIS>
 A>Note: the codons given for residues 78, 84, and 180 are inconsistent with the authors' sequence.
 R:Ratcliffe, M.; Emile, P.A.; Shannon, J.; Sano, K.; Hattler, B.; Mason, R.J.
 A:Title: Characterization of the rat pulmonary surfactant protein A promoter.
 A:Reference number: S23183; MUID:92298987; PMID:1606951
 A:Accession: S23183
 A:Molecule type: DNA
 A:Residues: 1-32 <LAC>
 A:Comment: Pulmonary surfactant is a complex of phospholipids and proteins that lowers the surface tension of the alveolar fluid.
 C:Comment: This protein is a salivary protein synthesized by alveolar type II cells. It is dependent on the presence of calcium ions.
 C:Comment: Two species of mRNA, which probably are transcribed from a single gene, have been identified.
 C:Superfamily: mannosyl-binding lectin; C-type lectin homology
 C:Keywords: alveolar proteinosis; calcium; gaseous exchange; glycoprotein; hydroxyproline
 F:1-20/Domain: signal sequence; status predicted <SIG>
 F:21-248/Product: pulmonary surfactant protein A; status experimental <MAT>
 F:37-109/Region: collagen-like
 F:127-246/Domain: C-type lectin homology <LCH>
 F:21/Binding site: carbohydrate (asn) (covalent) #status absent
 F:30,33,36,42,54,57,63,67,70,76/Modified site: 4-hydroxyproline (Pro) #status experimental
 F:207/Binding site: carbohydrate (asn) (covalent) #status predicted

Query Match 4.1%; Score 10; DB 1; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDG 54
 Db 40 GLPGRDGRDG 49

RESULT 14
 I51921
 pulmonary surfactant-associated protein A1 - human
 N:Alternate names: SP-A1
 C:Species: Homo sapiens (man)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 20-Aug-1999
 C:Accession: I51921
 R:Katyal, S.L.; Singh, G.; Locker, J.
 Am. J. Respir. Cell Mol. Biol. 6: 446-452, 1992
 A:Title: Characterization of a second human pulmonary surfactant-associated protein SP-A
 A:Reference number: I51921; MUID:92198680; PMID:1372511
 A:Accession: I51921
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-248 <RES>
 A:Cross-references: GB:M65519; NID:9338048; PID:AAA60319.1; PID:9338049
 C:Genetics:
 A:Gene: GDB:SFPA1; SFTPI; SP-A; SP-A1
 A:Cross-references: GDB:119593; OMIM:178630
 A:Map position: 10q22-10q23
 A:Introns: 58/1; 98/1; 124/1
 C:Superfamily: mannosyl-binding lectin; C-type lectin homology
 F:127-246/Domain: C-type lectin homology <LCH>

Query Match 4.1%; Score 10; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDG 54
 Db 40 GLPGRDGRDG 49

RESULT 15
 A48853
 pulmonary surfactant-associated protein SP-A - mouse
 C:Species: Mus musculus (house mouse)
 C:Date: 19-Dec-1993 #sequence_revision 18-Nov-1994 #text_change 20-Aug-1999
 C:Accession: A48853
 R:Korfhagen, T.R.; Bruno, M.D.; Glasser, S.W.; Ciruolo, P.J.; Whitsett, J.A.; Latliff, A.M.; J. Physiol. 263, 1546-1554, 1992
 A:Title: Murine pulmonary surfactant SP-A gene: cloning, sequence, and transcriptional regulation
 A:Reference number: A48853; MUID:93072386; PMID:1443158
 A:Accession: A48853
 A:Status: preliminary
 A:Molecule type: nucleic acid
 A:Residues: 1-248 <KOR>
 A:Cross-references: GB:S48768; NID:9260452; PID:AA824274.1; PID:9260453
 A:Note: Sequence extracted from NCBI backbone (NCBI:118740, NCBI:118741)
 C:Superfamily: mannosyl-binding lectin; C-type lectin homology
 F:127-246/Domain: C-type lectin homology <LCH>

Query Match 4.1%; Score 10; DB 2; Length 248;
 Best Local Similarity 100.0%; Pred. No. 0.13;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GLPGRDGRDG 54
 Db 40 GLPGRDGRDG 49

Search completed: June 18, 2003, 15:06:31
 Job time: 41 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 14:59:55 ; Search time 81 seconds

(without alignments)
618.141 Million cell updates/sec

Title: US-09-943-851A-42

Sequence: 1 MRPLVLLLLGLAGSPPLD.....DSTFGFLVYSDMHSSPVFA 243

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size : 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

1: SP archaea: *
2: SP bacteria: *
3: SP fungi: *
4: SP human: *
5: SP invertebrate: *
6: SP mammal: *
7: SP mhc: *
8: SP organelle: *
9: SP phage: *
10: SP plant: *
11: SP rodent: *
12: SP virus: *
13: SP vertebrate: *
14: SP unclassified: *
15: SP viirus: *
16: SP bacteriophage: *
17: SP archaea: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	25.5	243	11	08R002
2	11	4.5	27	4	09UCJ3
3	11	4.5	289	5	019813
4	11	4.5	306	2	09ACN2
5	11	4.5	375	4	08TCD8
6	11	4.5	378	6	09NIX4
7	10	4.1	57	6	002842
8	10	4.1	238	13	057431
9	10	4.1	248	13	091907
10	10	4.1	248	6	09T06
11	10	4.1	248	6	09N0G1
12	10	4.1	248	6	09SL88
13	10	4.1	248	11	09COI1
14	10	4.1	254	13	098T44
15	10	4.1	256	13	091909
16	10	4.1	282	5	08WP36

17	10	4.1	358	12	09J304	09J304 ectocarpus
18	10	4.1	447	12	09YTJ3	09YTJ3 atelline her
19	10	4.1	694	10	097T45	097T45 streptococc
20	10	4.1	812	5	006452	006452 ephydactia m
21	10	4.1	1761	5	018407	018407 dirosophila
22	10	4.1	1940	5	09YVW5	09YVW5 dirosophila
23	9	3.7	182	11	09CYSA	09CYSA mus musculu
24	9	3.7	260	4	09DHG2	09DHG2 mus saplen
25	9	3.7	260	11	09QXU9	09QXU9 ratius norv
26	9	3.7	441	13	09OYI9	09OYI9 brachydanio
27	9	3.7	567	10	09SMQ3	09SMQ3 artelmista a
28	9	3.7	589	16	09ARK4	09ARK4 streptomyce
29	9	3.7	684	5	P90679	P90679 arenlicola m
30	9	3.7	695	11	064348	064348 mus musculu
31	9	3.7	751	11	060709	060709 mus musculu
32	9	3.7	763	11	061482	061482 mus musculu
33	9	3.7	1752	5	007265	007265 strongyloce
34	9	3.7	2288	5	023081	023081 caenorhabdi
35	9	3.7	2315	5	095ZK3	095ZK3 caenorhabdi
36	8	3.3	46	11	063074	063074 ratius norv
37	8	3.3	52	16	09H209	09H209 pseudomonas
38	8	3.3	87	2	045313	045313 bacillus me
39	8	3.3	107	4	09BOY7	09BOY7 homo saplen
40	8	3.3	109	11	09CVJ2	09CVJ2 mus musculu
41	8	3.3	111	16	09RD99	09RD99 streptomyce
42	8	3.3	149	16	09JXB1	09JXB1 neisseria m
43	8	3.3	149	16	09JWV2	09JWV2 neisseria m
44	8	3.3	164	16	P74713	P74713 synchocyst
45	8	3.3	170	6	09GLV5	09GLV5 macaca mula

ALIGNMENTS

RESULT 1
ID 08R002 PRELIMINARY; PRT: 243 AA.
AC 08R002;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to DKFZ586B0621 protein (Hypothetical 25.4 kDa protein).
OS Mus musculus (Mouse).
CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (FEB-2002) to the EMBL/Genbank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/Genbank/DBJ databases.
DR EMBL: BC023068; AAH23068.1;
DR EMBL: BC025174; AAH25174.1;
KW Hypothetical Protein.
SQ SEQUENCE 243 AA; 25420 MW; 498129CD051DB97B CRC64;

Query Match 25.5%; Score 62; DB 11; Length 243;
Best Local Similarity 100.0%; Pred. No. 4.0e-52;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 FCGWKPASLSGAGAVRLPEPDQVWVGVDYIGIYASIKTDSFSGFLVYSDMHSSPV 241
DB 182 FCGWKPASLSGAGAVRLPEPDQVWVGVDYIGIYASIKTDSFSGFLVYSDMHSSPV 241
QY 242 FA 243
DB 242 FA 243

RESULT 2

09UCJ3
ID 09UCJ3 PRELIMINARY; PRT; 27 AA.
AC 09UCJ3;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DE 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE Surfactant protein D (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN NCBI_TaxID=9606;
RP
RX MEDLINE=93142849; PubMed=8424457;
RA Crouch E., Persson A., Chang D.;
RT "Accumulation of surfactant protein D in human pulmonary alveolar
RT proteinosis.";
RL Am. J. Pathol. 142:241-248(1993).
DR InterPro: IPR000087; Collagen
SQ SEQUENCE 27 AA; 2696 MW; 8E770650E2A967CA CRC64;

Query Match
Best Local Similarity 4.5%; Score 11; DB 4; Length 27;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRGR 55
DB 1 GLPGRDGRGR 11

RESULT 3

ID 019813 PRELIMINARY; PRT; 289 AA.
AC 019813;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE F26F12.1 protein.
GN F26F12.1.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodermidae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RX MEDLINE=39069613; PubMed=9851916;
RA None;
RT
RT "Genome sequence of the nematode C. elegans: a platform for
RT investigating biology. The C. elegans Sequencing Consortium.";
RL Science 282:2012-2018(1998).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Wilson R., Bentley D., Gattung S.;
RT "The sequence of C. elegans cosmid F26F12.";
RL Submitted (Apr-1996) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN-BRISTOL N2;
RA Waterston R.;
RL EMBL; U55373; AAC35888.1;
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR002486; Col_cuticle_N.
DR Pfam: PF01391; Collagen; 3.
DR Pfam: PF01484; Col_cuticle_N; 1.
SQ SEQUENCE 289 AA; 28740 MW; DFB78B2B34B29DF CRC64;

Query Match
Best Local Similarity 4.5%; Score 11; DB 5; Length 289;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 80 GPRGEGAPAG 90
DB 222 GPRGEGAPAG 232

RESULT 4

ID 09ACN2 PRELIMINARY; PRT; 306 AA.
AC 09ACN2;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE SC1B protein.
GN SC1B.
OS Streptococcus pyogenes.
OC Bacteria; Firmicutes; Bacilli; Clostridium group; Lactobacillales;
OC Streptococcaceae; Streptococcus.
OX NCBI_TaxID=1314;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=655;
RX MEDLINE=21097281; PubMed=11158359;
RA Whammar A.M.;
RT "Streptococcus pyogenes sc1B encodes a putative hypervariable surface
RT protein with a collagen-like repetitive structure.";
RL Microbiology 147:419-429(2001).
DR EMBL: AJ301809; CAC33778.1;
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001899; Gram_pos_anchor.
DR Pfam: PF00746; Collagen; 1.
DR Pfam: PF00746; Gram_pos_anchor; 1.
DR PRINTS: PR00015; GPOSANCHOR.
DR TIGR: TIGR01167; LPTXG_anchor; 1.
DR PROSITE: PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
KW Transmembrane.
SQ SEQUENCE 306 AA; 31833 MW; 57EBD244DF17600 CRC64;

Query Match
Best Local Similarity 4.5%; Score 11; DB 2; Length 306;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 81 GPRGEGAPAG 91
DB 148 GPRGEGAPAG 158

RESULT 5

ID 08TCD8 PRELIMINARY; PRT; 375 AA.
AC 08TCD8;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DE 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Hypothetical 37.7 Kda protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
RN NCBI_TaxID=9606;
RP
RX
RA TISSUE=LUNG;
RL Submitted (Feb-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC022318; AAH22318.1;
KW Hypothetical protein.
SQ SEQUENCE 375 AA; 37655 MW; CCB7375D3C86421A CRC64;

Query Match
Best Local Similarity 4.5%; Score 11; DB 4; Length 375;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRGR 55
DB 1 GLPGRDGRGR 11

DB 46 GLPGRDGRDGR 56

RESULT 6

ID 09N1X4 PRELIMINARY; PRT: 378 AA.

AC 09N1X4; 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

GN Lung surfactant protein D precursor.

OS Sus scrofa (pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.

OX NCBI_TaxID=9823;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=20109098; PubMed=10640760;

RA van Eljk M., Haegeman H.P., Skinner T., Archibald A.L., Reid K.B.M.,

RT Lawson P.R.;

RL Porcine Lung Surfactant Protein D (SP-D): cDNA cloning, chromosomal

J. Immunol. 164:1442-1450(2000).

DR EMBL: AF132496; AAF22145.2; -.

DR HSSP: P35247; 1808.

DR InterPro: IPR000087; Collagen.

DR InterPro: IPR001304; LECTIN_C.

DR Pfam: PF00059; Collagen; 3.

DR ProDom: PD000007; Collagen; 1.

DR SMART: SM00034; CLECT; 1.

DR PROSITE: PS00615; C-TYPE_LECTIN_1; 1.

DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.

FT SIGNAL.

FT CHAIN 1 20 POTENTIAL.

SQ SEQUENCE 378 AA; 37986 MW; 3504E8C1E56C341D CRC64;

Query Match

Best Local Similarity 4.58; Score 11; DB 6; Length 378;

Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDGR 55

DB 46 GLPGRDGRDGR 56

RESULT 7

ID 002842 PRELIMINARY; PRT: 57 AA.

AC 002842; 01-JUL-1997 (TREMblrel. 04, Created)

DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

OS Surfactant protein A2 (Fragment).

OC Papio cynocephalus (Yellow baboon).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Plimates; Catarrhini; Cercopithecidae;

OC Cercopithecinae; Papio.

OX NCBI_TaxID=9556;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=97053398; PubMed=8897910;

RA Gao E., Wang Y., McCormick S.M., Li J., Seidner S.R., Mendelson C.R.;

RT Characterization of two baboon surfactant protein A genes.*;

RL Am. J. Physiol. 271:L617-L630(1996).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE-KIDNEY;

RA Gao E., Wang Y., McCormick S.M., Li J., Seidner S.R., Mendelson C.R.;

RX Submitted (May-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF003897; AAB61295.1; -.

DR EMBL: AF003896; AAB61294.1; -.

DR InterPro: IPR000087; Collagen.

DR ProDom: PD000007; Collagen; 1.

FT NON_TER 57

SQ SEQUENCE 57 AA; 5584 MW; B25149EC822F643C CRC64;

Query Match

Best Local Similarity 4.18; Score 10; DB 6; Length 57;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDGR 54

DB 40 GLPGRDGRDGR 49

RESULT 8

ID 057451 PRELIMINARY; PRT: 238 AA.

AC 057451; 01-JUN-1998 (TREMblrel. 06, Created)

DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

GN Mannan-binding lectin (Fragment).

OS Gallus gallus (Chicken).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;

OC Gallus.

OX NCBI_TaxID=9031;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN-WHITE LEGHORN; TISSUE-LIVER;

RA Laursen S.B.;

RL Submitted (Aug-1997) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF022226; AAB94071.1; -.

DR HSSP: P19999; 1777.

DR InterPro: IPR000087; Collagen.

DR InterPro: IPR001304; LECTIN_C.

DR Pfam: PF00059; LECTIN_C; 1.

DR ProDom: PD000007; Collagen; 1.

DR SMART: SM00034; CLECT; 1.

DR PROSITE: PS00615; C-TYPE_LECTIN_1; UNKNOWN_1.

DR PROSITE: PS50041; C-TYPE_LECTIN_2; 1.

FT LECTIN.

FT NON_TER 238

SQ SEQUENCE 238 AA; 25645 MW; E5C9B5197AAB64E3 CRC64;

Query Match

Best Local Similarity 4.18; Score 10; DB 13; Length 238;

Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDGR 54

DB 31 GLPGRDGRDGR 40

RESULT 9

ID 091907 PRELIMINARY; PRT: 246 AA.

AC 091907; 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)

GN Mannose binding-like lectin precursor (Fragment).

OS Carassius auratus (Goldfish).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;

OC Cyprinidae; Carassius.

OX NCBI_TaxID=7957;

RN [1]

```

RP SEQUENCE FROM N.A.
RC TISSUE-LIVER;
RX MEDLINE-20456722; PubMed-11003389;
RA Vitved L., Holmskov U., Koch C., Teisner B., Hansen S., Skjoldt K.;
RT "The homologue of mannose-binding lectin in the carp family Cyprinidae
RT is expressed at high level in spleen, and the deduced primary
RL structure predicts affinity for galactose.";
DR EMBL: AF227739; AAF63470.1; -.
DR HSSP: P35247; 1B08.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF01391; Collagen; 2.
DR Pfam: PF00059; Lectin_C; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
DR LECTIN; Signal.
KM LECTIN; Signal.
FT MON_TER 1 1
FT SIGNAL <1 13 POTENTIAL.
FT VARIANT 145 145 S->F.
FT SEQUENCE 246 AA: 25709 MW: AB692282D289D0D5 CRC64;

Query Match 4.1%; Score 10; DB 13; Length 246;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDG 54
DB 35 GLPGRDGRDG 44

RESULT 10
O9TT06 PRELIMINARY; PRT; 248 AA.
AC O9TT06;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pulmonary surfactant protein A (Pulmonary surfactant-associated
DE protein A).
GN SPAS OR SP-A.
OS Ovis aries (Sheep).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-20215262; PubMed-10749753;
RA Braems G.A., Yao L.-J., Inchley K., Brickenden A., Han V.K.M.,
RA Grolla A., Challis J.R.G., Possmayer F.;
RT "Ovine surfactant protein CDNAS: use in studies on fetal lung growth
RT and maturation after prolonged hypoxemia.";
RL Am. J. Physiol. 278:L754-L764(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-LUNG;
RX Pletschmann S.M., Pison U.;
RT "cDNA cloning of ovine pulmonary SP-A, SP-B, and SP-C: Isolation of
RT two different sequences for SP-B.";
RL Am. J. Physiol. 278:L765-L778(2000).
DR EMBL: AF211856; AAF18995.1; -.
DR EMBL: AF076633; AAF31148.1; -.
DR HSSP: P23897; 1EGC.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF01391; Collagen; 2.
DR Pfam: PF00059; Lectin_C; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 248 AA: 26394 MW: D65E7293BBF1ED9 CRC64;

Query Match 4.1%; Score 10; DB 6; Length 248;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDG 54
DB 40 GLPGRDGRDG 49

RESULT 11
O9N0G1 PRELIMINARY; PRT; 248 AA.
AC O9N0G1;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pulmonary surfactant protein A.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=THOROUGHbred; TISSUE=LUNG;
RA Hodo S.;
RT "Molecular cloning of equine pulmonary surfactant proteins.";
RL Submitted (JUL-1998) to the EMBL/Genbank/DBJ databases.
DR EMBL: AB015963; BAA97976.1; -.
DR HSSP: P35247; 1B08.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF01391; Collagen; 2.
DR Pfam: PF00059; Lectin_C; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; 1.
DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 248 AA: 26000 MW: BB612FEFB05C2B8D1 CRC64;

Query Match 4.1%; Score 10; DB 6; Length 248;
Best Local Similarity 100.0%; Pred. No. 0.21;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDG 54
DB 40 GLPGRDGRDG 49

RESULT 12
O95L88 PRELIMINARY; PRT; 248 AA.
AC O95L88;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Pulmonary surfactant-associated protein A.
OS Equus caballus (Horse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Perissodactyla; Equidae; Equus.
OX NCBI_TaxID=9796;
RN [1]
RP SEQUENCE FROM N.A.
RA Weber B.L., Hospes R., Gortner L.;
RT Submitted (JUL-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF400580; AAL07690.1; -.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF01391; Collagen; 2.
DR Pfam: PF00059; Lectin_C; 1.
DR PROSITE: PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
DR PROSITE: PS50041; C_TYPE_LLECTIN_2; 1.
SQ SEQUENCE 248 AA: 26047 MW: B71133E005C9A5C1 CRC64;

```

Query Match
Best Local Similarity 4.1%; Score 10; DB 6; Length 248;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRCG 54
Db 40 GLPGRDGRCG 49

RESULT 13

09C011 PRELIMINARY; PRT; 248 AA.
AC 09C011: 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Surfactant associated protein A.
GN STPA.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]

SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=EMBRYO, AND LUNG;
RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Akawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Alizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,
RA Saito T., Okasaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barish G.,
RA Blake J., Boffelli D., Bojunga N., Carlini P., de Bonaldo M.F.,
RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustinich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,
RA Lyons P., Matchan L., Mashima J., Mazzarelli J., Mombearts P.,
RA Nodone P., Ring B., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Sasaki H., Sato K., Schoenbach C., Seta T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyszewski A., Yoshida K., Hasegawa Y., Kawai J., Kotsuki S.,
RA Hayashizaki Y.,
RT *Functional annotation of a full-length mouse cDNA collection.*;
RL Nature 409:685-690(2001).
DR EMBL: AK011333; BAB27551.1; -
DR EMBL: AK004620; BAB23416.1; -
DR EMBL: AK004788; BAB23565.1; -
DR HSSP: P35247; 1808.
DR MGD: MGI:109518; Stipa.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF01391; Collagen; 2.
DR Pfam: PF00059; Lectin_C; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; 1.
DR PROSITE: PS00041; C-TYPE_LLECTIN_2; 1.
SQ SEQUENCE 248 AA; 26183 MW; 8A5670CFAD3EB986 CRC64;

Query Match
Best Local Similarity 4.1%; Score 10; DB 11; Length 248;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRCG 54
Db 40 GLPGRDGRCG 49

RESULT 14
098TA4 PRELIMINARY; PRT; 254 AA.
ID 098TA4

AC 098TA4:
DT 01-JUN-2001 (TREMblrel. 17, Created)
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Mannose-binding lectin protein precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]

SEQUENCE FROM N.A.
RX MEDLINE=20456722; PubMed=11003389;
RA Vilved L., Holmskov U., Koch C., Telsner B., Hansen S., Skjold K.;
RT *The homologue of mannose-binding lectin in the carp family Cyprinidae
is expressed at high level in spleen, and the deduced primary
structure predicts affinity for galactose.*;
RL Immunogenetics 51:955-964(2000).

DR EMBL: AF21714; AAK30298.1; -
DR HSSP: P19999; 1YTT.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR000005; HTHAAC.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF01391; Collagen; 1.
DR Pfam: PF00059; Lectin_C; 1.
DR ProDom: PD000007; Collagen; 1.
DR SMART: SM00034; CLECT; 1.
DR PROSITE: PS00615; C-TYPE_LLECTIN_1; UNKNOWN_1.
DR PROSITE: PS00041; C-TYPE_LLECTIN_2; 1.
DR PROSITE: PS00041; HTH_ARAC_FAMILY_1; UNKNOWN_1.
KW Lectin; Signal.

FT SIGNAL 1 21 POTENTIAL.
FT CHAIN 22 254 MANNOSE-BINDING LECTIN PROTEIN.
FT VARIANT 234 234 D -> V.
SQ SEQUENCE 254 AA; 27376 MW; C924428643441AED CRC64;

Query Match
Best Local Similarity 4.1%; Score 10; DB 13; Length 254;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRCG 54
Db 47 GLPGRDGRCG 56

RESULT 15
091909 PRELIMINARY; PRT; 256 AA.
ID 091909

AC 091909: 01-OCT-2000 (TREMblrel. 15, Created)
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMblrel. 21, Last annotation update)
DE Mannose binding-like lectin precursor.
GN MBL.

OS Cyprinus carpio (Common carp).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Cyprinus.
OX NCBI_TaxID=7962;
RN [1]

SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RX MEDLINE=20456722; PubMed=11003389;

RA Vilved L., Holmskov U., Koch C., Telsner B., Hansen S., Skjold K.;
RT *The homologue of mannose-binding lectin in the carp family Cyprinidae
is expressed at high level in spleen, and the deduced primary
structure predicts affinity for galactose.*;
RL Immunogenetics 51:955-964(2000).

DR EMBL: AF227737; AAF63468.1; -
DR HSSP: P35247; 1808.
DR InterPro: IPR000087; Collagen.
DR InterPro: IPR001304; Lectin_C.
DR Pfam: PF01391; Collagen; 2.

DR Pfam; PF00059; lectin_c; 1.
 DR SMART; SM00034; CLECT; 1.
 DR PROSITE; PS00615; C_TYPE_LLECTIN_1; UNKNOWN_1.
 DR PROSITE; PS50041; C_TYPE_LLECTIN_2; 1.
 KW lectin; signal.
 FT SIGNAL 1 23
 FT CHAIN 24 256
 FT VARIANT 235 235
 SQ SEQUENCE 256 AA; 26934 MW; D019291D1167730D CRC64;

Query Match 4.1%; Score 10; DB 13; Length 256;
 Best Local Similarity 100.0%; Pred. No. 0.22;
 Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 45 GLPGRDGRDG 54
 DB 45 GLPGRDGRDG 54

Search completed: June 18, 2003, 15:05:52
 Job time : 91 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:02:44 ; Search time 26 seconds

(Without alignments)
274.991 Million cell updates/sec

Title: US-09-943-851a-42

Perfect score: 243
Sequence: 1 MRPLVLLGLAAGSPPLD.....DSTFGFLVYSMDHSSPVFA 243

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database : Issued Patents_AA:*

1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:.*
2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:.*
3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:.*
4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:.*
5: /cgn2_6/ptodata/1/1aa/PCPUS.COMB.pep:.*
6: /cgn2_6/ptodata/1/1aa/backfiles1.pep:.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	243	100.0	243	US-09-140-804-2	Sequence 2, Appl1
2	243	100.0	243	US-09-336-536-3	Sequence 3, Appl1
3	228	93.8	228	US-09-336-536-4	Sequence 4, Appl1
4	128	52.7	128	US-09-336-536-7	Sequence 7, Appl1
5	70	28.8	243	US-09-188-930-295	Sequence 295, App
6	62	25.5	228	US-09-336-536-11	Sequence 11, Appl1
7	62	25.5	243	US-09-336-536-10	Sequence 10, Appl1
8	60	24.7	60	US-09-336-536-6	Sequence 6, Appl1
9	51	21.0	128	US-09-336-536-14	Sequence 14, Appl1
10	43	17.7	60	US-09-336-536-13	Sequence 13, Appl1
11	15	6.2	15	US-09-336-536-5	Sequence 5, Appl1
12	11	4.5	128	US-09-227-357-190	Sequence 190, App
13	10	4.1	1057	US-08-931-820-4	Sequence 4, Appl1
14	10	4.1	1078	US-08-963-825-21	Sequence 21, Appl1
15	10	4.1	1078	US-09-500-811-21	Sequence 21, Appl1
16	10	4.1	1078	US-09-570-573-21	Sequence 21, Appl1
17	10	4.1	1078	US-09-548-608-21	Sequence 21, Appl1
18	10	4.1	1078	US-09-548-608-21	Sequence 21, Appl1
19	8	3.3	130	US-09-485-316A-10	Sequence 10, Appl1
20	8	3.3	170	US-08-313-681A-2	Sequence 2, Appl1
21	8	3.3	170	US-09-322-911-2	Sequence 2, Appl1
22	8	3.3	184	US-08-211-942-7	Sequence 7, Appl1
23	8	3.3	184	US-08-211-942-9	Sequence 9, Appl1
24	8	3.3	222	US-09-140-804-7	Sequence 7, Appl1
25	8	3.3	231	US-09-530-423-2	Sequence 2, Appl1
26	8	3.3	244	US-08-463-911-7	Sequence 7, Appl1
27	8	3.3	244	US-09-140-804-3	Sequence 3, Appl1

28	8	3.3	244	US-09-336-536-20	Sequence 20, Appl1
29	8	3.3	247	US-09-530-423-1	Sequence 1, Appl1
30	8	3.3	247	US-08-463-911-2	Sequence 2, Appl1
31	8	3.3	247	US-09-140-804-8	Sequence 8, Appl1
32	8	3.3	247	US-09-118-408-3	Sequence 3, Appl1
33	8	3.3	247	US-09-506-855-3	Sequence 3, Appl1
34	8	3.3	368	US-08-211-942-17	Sequence 17, Appl1
35	8	3.3	547	US-08-494-168-7	Sequence 7, Appl1
36	8	3.3	1024	US-08-931-820-2	Sequence 2, Appl1
37	8	3.3	1366	US-08-963-825-19	Sequence 19, Appl1
38	8	3.3	1366	US-09-500-811-19	Sequence 19, Appl1
39	8	3.3	1366	US-09-570-573-19	Sequence 19, Appl1
40	8	3.3	1366	US-09-548-608-19	Sequence 19, Appl1
41	8	3.3	1366	US-09-585-887-10	Sequence 10, Appl1
42	8	3.3	1366	US-09-289-578-10	Sequence 10, Appl1
43	8	3.3	1442	US-08-316-650-12	Sequence 12, Appl1
44	8	3.3	1442	PCT-US95-02251-12	Sequence 12, Appl1
45	8	3.3	1694	US-08-494-168-2	Sequence 2, Appl1

ALIGNMENTS

RESULT 1
US-09-140-804-2
; Sequence 2, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-49
; CURRENT APPLICATION NUMBER: US/09/140, 804
; EARLIER FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/056, 983
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-140-804-2

Query Match 100.0%; Score 243; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-218;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MRPLVLLGLAAGSPPLDNNKIPSLCPGHPCLPCTPGHSGQLPGRDGRDGRGARG	60
DB	1	MRPLVLLGLAAGSPPLDNNKIPSLCPGHPCLPCTPGHSGQLPGRDGRDGRGARG	60
QY	61	APPEKGGGRGRLPGRCGDPGRGEAGAPPTGAGCSVPPRSARSRSPRPSPD	120
DB	61	APPEKGGGRGRLPGRCGDPGRGEAGAPPTGAGCSVPPRSARSRSPRPSPD	120
QY	121	APLPRLVNEGCHDAVTKGFTQVPCVYFAVAHATYRASLDFDLVKNESIASFPO	180
DB	121	APLPRLVNEGCHDAVTKGFTQVPCVYFAVAHATYRASLDFDLVKNESIASFPO	180
QY	181	FFGMPKPPASLSCGAVRLEPEDQVWVGVDYIGIVASITDSTFGFLVYSMDHSSP	240
DB	181	FFGMPKPPASLSCGAVRLEPEDQVWVGVDYIGIVASITDSTFGFLVYSMDHSSP	240
QY	241	VFA 243	
DB	241	VFA 243	

RESULT 2
US-09-336-536-3
; Sequence 3, Application US/09336536
; Patent No. 6406884

GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
NUMBER OF SEQ ID NOS: 75
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 3
LENGTH: 243
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-3

Query Match 100.0%; Score 243; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-218;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGRDAPG 60
DB 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGRDAPG 60
QY 61 APEGEGEGRPLPGPRGDPGPRGAPGPTGAGGECVPPRSASFSAKRSSESVPPSD 120
DB 61 APEGEGEGRPLPGPRGDPGPRGAPGPTGAGGECVPPRSASFSAKRSSESVPPSD 120
QY 121 ALPLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLOPDLVKNGESIASFFQ 180
DB 121 ALPLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLOPDLVKNGESIASFFQ 180
QY 181 FFGGMPKPSLSCGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFGFLVSDMHSSP 240
DB 181 FFGGMPKPSLSCGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFGFLVSDMHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 3
US-09-336-536-4
Sequence 4, Application US/09336536
Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
NUMBER OF SEQ ID NOS: 75
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 4
LENGTH: 228
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-4

Query Match 93.8%; Score 228; DB 4; Length 228;
Best Local Similarity 100.0%; Pred. No. 1.4e-204;
Matches 228; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 16 SPPLDDNKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGRDAPGAGGEGGRPLPG 75
DB 1 SPPLDDNKIPSLCPGHPGLPTGPHHGSQGLPGRDGRDGRDAPGAGGEGGRPLPG 75
QY 76 PRGDPGPRGAPGPTGAGGECVPPRSASFSAKRSSESVPPSDALPLPDRVLVNEQGH 135
DB 61 PRGDPGPRGAPGPTGAGGECVPPRSASFSAKRSSESVPPSDALPLPDRVLVNEQGH 120

QY 136 YDAVTGKFTCOVPGVYFAVHATVYRASLOPDLVKNGESIASFFQFGGMPKPSLSCG 195
DB 121 YDAVTGKFTCOVPGVYFAVHATVYRASLOPDLVKNGESIASFFQFGGMPKPSLSCG 180
QY 196 MVRLPEPDQVWVGVGDYIGIYASIKTDSFGFLVSDMHSSPVEA 243
DB 181 MVRLPEPDQVWVGVGDYIGIYASIKTDSFGFLVSDMHSSPVEA 228

RESULT 4
US-09-336-536-7
Sequence 7, Application US/09336536
Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: Leiby, K.
APPLICANT: McKay, C.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
NUMBER OF SEQ ID NOS: 75
SOFTWARE: Patentin Ver. 2.0
SEQ ID NO: 7
LENGTH: 128
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-7

Query Match 52.7%; Score 128; DB 4; Length 128;
Best Local Similarity 100.0%; Pred. No. 9.7e-112;
Matches 128; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 105 AFSAKRSSESVPPSDALPLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASL 164
DB 1 AFSAKRSSESVPPSDALPLPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASL 60
QY 165 QFDLVKNGESIASFFQFGGMPKPSLSCGAMVRLPEPDQVWVGVGDYIGIYASIKTD 224
DB 61 QFDLVKNGESIASFFQFGGMPKPSLSCGAMVRLPEPDQVWVGVGDYIGIYASIKTD 120
QY 225 STESGFLV 232
DB 121 STESGFLV 128

RESULT 5
US-09-188-930-295
Sequence 295, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Muriison, James Greg
TITLE OF INVENTION: Compositions Isolated from Skin Cells
FILE REFERENCE: 11000.1011c1
CURRENT APPLICATION NUMBER: US/09/188,930A
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO: 295
LENGTH: 243
TYPE: PRT
ORGANISM: Rat
US-09-188-930-295

Query Match 28.8%; Score 70; DB 4; Length 243;
Best Local Similarity 100.0%; Pred. No. 1.6e-57;
Matches 70; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 174 SIASFQFGGMPKASLSGAMVRLPEPDQVWVGVGDYIGIYASIKTSTFSGFLVY 233
DB 174 SIASFQFGGMPKASLSGAMVRLPEPDQVWVGVGDYIGIYASIKTSTFSGFLVY 233
OY 234 SDMHSSPVFA 243
DB 234 SDMHSSPVFA 243

RESULT 6
US-09-336-536-11
; Sequence 11, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11.
; LENGTH: 228
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-336-536-11

Query Match
Best Local Similarity 100.0%; Pred. No. 4.2e-50;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 182 FCGMPKASLSGAMVRLPEPDQVWVGVGDYIGIYASIKTSTFSGFLVYSDMHSSPV 241
DB 167 FCGMPKASLSGAMVRLPEPDQVWVGVGDYIGIYASIKTSTFSGFLVYSDMHSSPV 226
OY 242 FA 243
DB 227 FA 228

RESULT 7
US-09-336-536-10
; Sequence 10, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-336-536-10

Query Match
Best Local Similarity 100.0%; Pred. No. 4.5e-50;
Matches 62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 182 FCGMPKASLSGAMVRLPEPDQVWVGVGDYIGIYASIKTSTFSGFLVYSDMHSSPV 241
DB 182 FCGMPKASLSGAMVRLPEPDQVWVGVGDYIGIYASIKTSTFSGFLVYSDMHSSPV 241
OY 242 FA 243
DB 242 FA 243

RESULT 8
US-09-336-536-6
; Sequence 6, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-336-536-6

Query Match
Best Local Similarity 100.0%; Pred. No. 9.6e-49;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 36 GPGHGSQGLPGRDGRDGPAGPAGKEGEGRPGLPGRDDPGPRGAPGAGPTGPA 95
DB 1 GPGHGSQGLPGRDGRDGPAGPAGKEGEGRPGLPGRDDPGPRGAPGAGPTGPA 60

RESULT 9
US-09-336-536-14
; Sequence 14, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14
; LENGTH: 128
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-336-536-14

Query Match
Best Local Similarity 100.0%; Pred. No. 4.4e-40;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 182 FCGMPKASLSGAMVRLPEPDQVWVGVGDYIGIYASIKTSTFSGFLVY 232
DB 78 FCGMPKASLSGAMVRLPEPDQVWVGVGDYIGIYASIKTSTFSGFLVY 128

RESULT 10
US-09-336-536-13
; Sequence 13, Application US/09336536
; Patent No. 6406884
; GENERAL INFORMATION:
; APPLICANT: Leiby, K.
; APPLICANT: McKay, C.
; APPLICANT: Bossone, S.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-144
; CURRENT APPLICATION NUMBER: US/09/336,536
; CURRENT FILING DATE: 1999-06-18
; NUMBER OF SEQ ID NOS: 75

SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 13
LENGTH: 60
TYPE: PRT
ORGANISM: Mus musculus
US-09-336-536-13

Query Match
Best Local Similarity 17.7%; Score 43; DB 4; Length 60;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 36 GTPGHGSGQLPGRDGRGARGGKGGGRLPPRG 78
DB 1 GTPGHGSGQLPGRDGRGARGGKGGGRLPPRG 43

RESULT 11
US-09-336-536-5
Sequence 5, Application US/09336536
Patent No. 6406884
GENERAL INFORMATION:
APPLICANT: McKay, K.
APPLICANT: Bosson, S.
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
FILE REFERENCE: 7853-144
CURRENT APPLICATION NUMBER: US/09/336,536
CURRENT FILING DATE: 1999-06-18
NUMBER OF SEQ ID NOS: 75
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 15
TYPE: PRT
ORGANISM: Homo sapiens
US-09-336-536-5

Query Match
Best Local Similarity 6.2%; Score 15; DB 4; Length 15;
Matches 15; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAG 15
DB 1 MRPLVLLLLGLAAG 15

RESULT 12
US-09-227-357-190
Sequence 190, Application US/09227357
Patent No. 6342581
GENERAL INFORMATION:
APPLICANT: Fischer et al.
TITLE OF INVENTION: 123 Human Secreted Proteins
FILE REFERENCE: P2010P1
CURRENT APPLICATION NUMBER: US/09/227,357
EARLIER APPLICATION NUMBER: 1999-01-08
EARLIER FILING DATE: 1998-07-07
EARLIER APPLICATION NUMBER: 60/051,926
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,793
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,925
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,929
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,803
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,732
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,931
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,932
EARLIER FILING DATE: 1997-07-08

EARLIER APPLICATION NUMBER: 60/051,916
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,930
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,918
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,920
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,733
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/052,795
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,919
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/051,928
EARLIER FILING DATE: 1997-07-08
EARLIER APPLICATION NUMBER: 60/055,722
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,723
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,948
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,949
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,953
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,950
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,947
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,964
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/056,360
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,684
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,984
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/055,954
EARLIER FILING DATE: 1997-08-18
EARLIER APPLICATION NUMBER: 60/058,785
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,664
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,660
EARLIER FILING DATE: 1997-09-12
EARLIER APPLICATION NUMBER: 60/058,661
EARLIER FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 672
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 190
LENGTH: 128
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SITE
LOCATION: (127)
OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
NAME/KEY: SITE
LOCATION: (128)
OTHER INFORMATION: Xaa equals stop translation
US-09-227-357-190

Query Match
Best Local Similarity 4.5%; Score 11; DB 4; Length 128;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 45 GTPGRDGR 55
DB 59 GTPGRDGR 69

RESULT 13
US-08-931-820-4
Sequence 4, Application US/08931820
Patent No. 6010863
GENERAL INFORMATION:
APPLICANT:
TITLE OF INVENTION: Assay for collagen degradation
NUMBER OF SEQUENCES: 4
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25 (PPO)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/931,820
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: EP 96202596.1
FILING DATE:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 1057 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
HYPOTHECAL: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapiens
TISSUE TYPE: Collagen type III
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1055
OTHER INFORMATION: /label= Modified
OTHER INFORMATION: /note= Ala may be pro
US-08-931-820-4

Query Match
Best Local Similarity 4.1%; Score 10; DB 3; Length 1057;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 GAGPKGEGC 69
DB 680 GAGPKGEGC 689

RESULT 14
US-08-963-825-21
Sequence 21, Application US/08963825
Patent No. 6110689
GENERAL INFORMATION:
APPLICANT: Ovist, Per
TITLE OF INVENTION: A Method for Assaying Collagen Fragments
TITLE OF INVENTION: In Body Fluids, A Test Kit and Means for Carrying Out the
TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/963,825

FILING DATE: 436
CLASSIFICATION: 436
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/187,319
FILING DATE: 21-JAN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda C
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 4305/08701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687
US-08-963-825-21

Query Match
Best Local Similarity 4.1%; Score 10; DB 3; Length 1078;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 60 GAGPKGEGC 69
DB 681 GAGPKGEGC 690

RESULT 15
US-09-500-811-21
Sequence 21, Application US/09500811
Patent No. 6323314
GENERAL INFORMATION:
APPLICANT: Ovist, Per
TITLE OF INVENTION: A Method for Assaying Collagen Fragments
TITLE OF INVENTION: In Body Fluids, A Test Kit and Means for Carrying Out the
TITLE OF INVENTION: Method and Use of the Method to Diagnose the Presence of
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Darby & Darby PC
STREET: 805 Third Avenue
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/500,811
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/187,319
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Gogoris, Adda C
REGISTRATION NUMBER: 29,714
REFERENCE/DOCKET NUMBER: 4305/08701
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-527-7700
TELEFAX: 212-753-6237
TELEX: 236687

```

; INFORMATION FOR SEQ ID NO: 21:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
; IMMEDIATE SOURCE:
; CLONE: COLLAGEN ALPHA 1 (III)
US-09-500-811-21

```

```

Query Match      4.18; Score 10; DB 4; Length 1078;
Best Local Similarity 100.0%; Pred. No. 0.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      60 GAPGEGEGG 69
        |||
Db      681 GAPGEGEGG 690

```

Search completed: June 18, 2003, 15:07:04
 Job time : 27 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 18, 2003, 15:05:58 ; Search time 49 Seconds
(Without alignments)
536.617 Million cell updates/sec

Title: US-09-943-851a-42
Sequence: 1 MRPLVLLGLAGSPPLD.....DSFSGFLVYSDHSSPVFA 243

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 41779 seqs, 108206813 residues

Word size : 0

Total number of hits satisfying chosen parameters: 41779

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/2/pubppa/US08_NEW_PUB.pep.*
2: /cgn2_6/ptodata/2/pubppa/PTCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/2/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/2/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/2/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/2/pubppa/US07_PUBCOMB.pep.*
7: /cgn2_6/ptodata/2/pubppa/PTCTUS_PUBCOMB.pep.*
8: /cgn2_6/ptodata/2/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/2/pubppa/US09_NEW_PUB.pep.*
10: /cgn2_6/ptodata/2/pubppa/US09_PUBCOMB.pep.*
11: /cgn2_6/ptodata/2/pubppa/US10_NEW_PUB.pep.*
12: /cgn2_6/ptodata/2/pubppa/US10_PUBCOMB.pep.*
13: /cgn2_6/ptodata/2/pubppa/US60_NEW_PUB.pep.*
14: /cgn2_6/ptodata/2/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	243	100.0	243	9	US-09-944-413-42
2	243	100.0	243	9	US-09-944-403-42
3	243	100.0	243	9	US-09-944-896-42
4	243	100.0	243	9	US-09-944-944-42
5	243	100.0	243	9	US-09-944-907-42
6	243	100.0	243	9	US-09-944-929-42
7	243	100.0	243	9	US-10-028-072-362
8	243	100.0	243	9	US-10-121-049-362
9	243	100.0	243	9	US-10-123-904-362
10	243	100.0	243	9	US-10-140-470-362
11	243	100.0	243	9	US-09-796-753-68
12	243	100.0	243	9	US-10-175-746-362
13	243	100.0	243	9	US-10-176-918-362
14	243	100.0	243	9	US-10-176-921-362
15	243	100.0	243	9	US-10-137-865-362
16	243	100.0	243	9	US-10-140-474-362
17	243	100.0	243	9	US-10-143-431-362
18	243	100.0	243	9	US-10-143-114-362
19	243	100.0	243	9	US-10-140-002-362

20	243	100.0	243	9	US-10-147-419-362	Sequence 362, App
21	243	100.0	243	9	US-10-123-262-362	Sequence 362, App
22	243	100.0	243	9	US-10-142-423-362	Sequence 362, App
23	243	100.0	243	9	US-10-121-050-362	Sequence 362, App
24	243	100.0	243	9	US-10-141-755-362	Sequence 362, App
25	243	100.0	243	9	US-10-143-032-362	Sequence 362, App
26	243	100.0	243	9	US-10-123-108-362	Sequence 362, App
27	243	100.0	243	9	US-10-123-261-362	Sequence 362, App
28	243	100.0	243	9	US-10-140-928-362	Sequence 362, App
29	243	100.0	243	9	US-10-121-045-362	Sequence 362, App
30	243	100.0	243	9	US-10-123-292-362	Sequence 362, App
31	243	100.0	243	9	US-10-123-903-362	Sequence 362, App
32	243	100.0	243	9	US-10-124-819-362	Sequence 362, App
33	243	100.0	243	9	US-10-124-822-362	Sequence 362, App
34	243	100.0	243	9	US-10-140-923-362	Sequence 362, App
35	243	100.0	243	9	US-09-944-884-42	Sequence 42, App
36	243	100.0	243	9	US-10-121-041-362	Sequence 362, App
37	243	100.0	243	9	US-10-121-047-362	Sequence 362, App
38	243	100.0	243	9	US-10-123-215-362	Sequence 362, App
39	243	100.0	243	9	US-10-123-908-362	Sequence 362, App
40	243	100.0	243	9	US-10-123-908-362	Sequence 362, App
41	243	100.0	243	9	US-10-123-908-362	Sequence 362, App
42	243	100.0	243	9	US-10-123-908-362	Sequence 362, App
43	243	100.0	243	9	US-10-123-908-362	Sequence 362, App
44	243	100.0	243	9	US-10-123-908-362	Sequence 362, App
45	243	100.0	243	9	US-10-123-908-362	Sequence 362, App

ALIGNMENTS

RESULT 1
US-09-944-413-42
Sequence 42, Application US/09944413
Patent No. US200201560041
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gertlisen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kljavin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Thomas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,413
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 60/067,411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,686
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997

;; PRIOR APPLICATION NUMBER: 60/069,702
;; PRIOR FILING DATE: December 16, 1997
;; PRIOR APPLICATION NUMBER: 60/069,870
;; PRIOR FILING DATE: December 17, 1997
;; PRIOR APPLICATION NUMBER: 60/069,873
;; PRIOR FILING DATE: December 17, 1997
;; PRIOR APPLICATION NUMBER: 60/068,017
;; PRIOR FILING DATE: December 18, 1997
;; PRIOR APPLICATION NUMBER: 60/070,440
;; PRIOR FILING DATE: January 5, 1998
;; PRIOR APPLICATION NUMBER: 60/074,086
;; PRIOR FILING DATE: February 9, 1998
;; PRIOR APPLICATION NUMBER: 60/074,092
;; PRIOR FILING DATE: February 9, 1998
;; PRIOR APPLICATION NUMBER: 60/075,945
;; PRIOR FILING DATE: February 25, 1998
;; PRIOR APPLICATION NUMBER: 60/112,850
;; PRIOR FILING DATE: December 16, 1998
;; PRIOR APPLICATION NUMBER: 60/113,296
;; PRIOR FILING DATE: December 22, 1998
;; PRIOR APPLICATION NUMBER: 60/146,222
;; PRIOR FILING DATE: July 28, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/19330
;; PRIOR FILING DATE: September 16, 1998
;; PRIOR APPLICATION NUMBER: PCT/US98/25108
;; PRIOR FILING DATE: December 1, 1998
;; PRIOR APPLICATION NUMBER: 09/216,021
;; PRIOR FILING DATE: December 16, 1998
;; PRIOR APPLICATION NUMBER: 09/218,517
;; PRIOR FILING DATE: December 22, 1998
;; PRIOR APPLICATION NUMBER: 09/254,311
;; PRIOR FILING DATE: March 3, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/12252
;; PRIOR FILING DATE: June 22, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/21090
;; PRIOR FILING DATE: September 15, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/28409
;; PRIOR FILING DATE: NO. US20020156004A1ember 30, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: NO. US20020156004A1ember 30, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/28301
;; PRIOR FILING DATE: December 1, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: December 16, 1999
;; PRIOR APPLICATION NUMBER: PCT/US00/03565
;; PRIOR FILING DATE: February 11, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/04414
;; PRIOR FILING DATE: February 22, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: March 2, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/08439
;; PRIOR FILING DATE: March 30, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/14042
;; PRIOR FILING DATE: May 22, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/20710
;; PRIOR FILING DATE: July 28, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: December 1, 2000
;; PRIOR APPLICATION NUMBER: PCT/US01/06520
;; PRIOR FILING DATE: February 28, 2001
;; NUMBER OF SEQ ID NOS: 120
;; SEQ ID NO 42
;; LENGTH: 243
;; TYPE: PRT
;; ORGANISM: Homo Saplen
US-09-944-413-42

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Freq. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHSGQLPGRDGRDGRDAGP 60
|||||

Db 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHSGQLPGRDGRDGRDAGP 60
Qy 61 AGECEGEGRCPELPGRDPPRGEAGPAGPTGAGCCSVPRSAFSAKSESVPSPD 120
Db 61 AGECEGEGRCPELPGRDPPRGEAGPAGPTGAGCCSVPRSAFSAKSESVPSPD 120
Qy 121 ALPFDRLVNEOGHYDAVTKFTCOVPGVYFVAVATVYRASLQFDLVNKGESIASFFQ 180
Db 121 ALPFDRLVNEOGHYDAVTKFTCOVPGVYFVAVATVYRASLQFDLVNKGESIASFFQ 180
Qy 181 FFGGPKRSLSSGAMVLEPEDQVWVQVGVGYIGIYASIKTDSFSGFLVYSMDHSSP 240
Db 181 FFGGPKRSLSSGAMVLEPEDQVWVQVGVGYIGIYASIKTDSFSGFLVYSMDHSSP 240
Qy 241 VFA 243
Db 241 VFA 243

RESULT 2
US-09-944-403-42
;; Sequence 42, Application US/09944403
;; Patent No. US20020165143A1
;; GENERAL INFORMATION:
;; APPLICANT: Baker, Kevin
;; APPLICANT: Botstein, David
;; APPLICANT: Eaton, Dan
;; APPLICANT: Ferrara, Napoleone
;; APPLICANT: Filvaroff, Ellen
;; APPLICANT: Gertsen, Mary
;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul
;; APPLICANT: Grimaldi, Christopher
;; APPLICANT: Gurney, Austin
;; APPLICANT: Hillan, Kenneth
;; APPLICANT: Kijavlin, Ivar
;; APPLICANT: Napier, Mary
;; APPLICANT: Roy, Margaret
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Wood, William
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P2348PICI
;; CURRENT FILING DATE: US/09/944,403
;; PRIOR APPLICATION NUMBER: 2001-09-26
;; PRIOR FILING DATE: 2001-05-25
;; PRIOR APPLICATION NUMBER: 60/067,411
;; PRIOR FILING DATE: December 3, 1997
;; PRIOR APPLICATION NUMBER: 60/069,334
;; PRIOR FILING DATE: December 11, 1997
;; PRIOR APPLICATION NUMBER: 60/069,335
;; PRIOR FILING DATE: December 11, 1997
;; PRIOR APPLICATION NUMBER: 60/069,278
;; PRIOR FILING DATE: December 11, 1997
;; PRIOR APPLICATION NUMBER: 60/069,425
;; PRIOR FILING DATE: December 12, 1997
;; PRIOR APPLICATION NUMBER: 60/069,696
;; PRIOR FILING DATE: December 15, 1997
;; PRIOR APPLICATION NUMBER: 60/069,694
;; PRIOR FILING DATE: December 16, 1997
;; PRIOR APPLICATION NUMBER: 60/069,702
;; PRIOR FILING DATE: December 16, 1997
;; PRIOR APPLICATION NUMBER: 60/069,870
;; PRIOR FILING DATE: December 17, 1997
;; PRIOR APPLICATION NUMBER: 60/069,873
;; PRIOR FILING DATE: December 17, 1997
;; PRIOR APPLICATION NUMBER: 60/068,017
;; PRIOR FILING DATE: December 18, 1997
;; PRIOR APPLICATION NUMBER: 60/070,440
;; PRIOR FILING DATE: January 5, 1998
;; PRIOR APPLICATION NUMBER: 60/074,086
;; PRIOR FILING DATE: February 9, 1998

PRIOR APPLICATION NUMBER: 60/074,092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075,945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112,850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113,296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146,222
PRIOR FILING DATE: July 28, 1998
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998
PRIOR APPLICATION NUMBER: 09/216,021
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 09/218,517
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 09/254,311
PRIOR FILING DATE: March 3, 1999
PRIOR APPLICATION NUMBER: PCT/US99/12252
PRIOR FILING DATE: June 22, 1999
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: September 15, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28409
PRIOR FILING DATE: No. US20020165143A1eember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: No. US20020165143A1eember 30, 1999
PRIOR APPLICATION NUMBER: PCT/US99/28301
PRIOR FILING DATE: December 1, 1999
PRIOR APPLICATION NUMBER: PCT/US99/30005
PRIOR FILING DATE: December 16, 1999
PRIOR APPLICATION NUMBER: PCT/US00/03365
PRIOR FILING DATE: February 11, 2000
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: February 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: March 2, 2000
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 42
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-09-944-403-42

Query Match 100.0% Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Caps 0;

1 MRLVLLVLLGLAGSPPLDNKIPSLCPGHPGLPGTRGHHSQGLPGRGRGRCAGC 60
1 MRLVLLVLLGLAGSPPLDNKIPSLCPGHPGLPGTRGHHSQGLPGRGRGRCAGC 60
61 APGEKGGGRRGLGPPGDDPGRGECAPACPTGACGSCVPPRSASAKRSSESRVPPSD 120
61 APGEKGGGRRGLGPPGDDPGRGECAPACPTGACGSCVPPRSASAKRSSESRVPPSD 120
121 APLPFDRLVNEGQHYAVTGKTCOVFGVYFAVAHATYRASLOPDLVKNESIASFFQ 180
121 APLPFDRLVNEGQHYAVTGKTCOVFGVYFAVAHATYRASLOPDLVKNESIASFFQ 180
121 APLPFDRLVNEGQHYAVTGKTCOVFGVYFAVAHATYRASLOPDLVKNESIASFFQ 180
181 FFGGMPKASISGAGWRLPEPDQVWVGVDYIGIYASITDSTFGSLVSDMHSSP 240
181 FFGGMPKASISGAGWRLPEPDQVWVGVDYIGIYASITDSTFGSLVSDMHSSP 240

DB 181 FFGGMPKASISGAGWRLPEPDQVWVGVDYIGIYASITDSTFGSLVSDMHSSP 240
QY 241 VEA 243
DB 241 VEA 243
RESULT 3
US-09-944-896-42
Sequence 42; Application US/09944896
Patent No. US20020168713A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Balon, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerlitsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tamas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
TITLE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944, 896
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866, 028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/069, 334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069, 425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069, 696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069, 694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069, 702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069, 870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069, 873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068, 017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070, 440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074, 086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074, 092
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/075, 945
PRIOR FILING DATE: February 25, 1998
PRIOR APPLICATION NUMBER: 60/112, 850
PRIOR FILING DATE: December 16, 1998
PRIOR APPLICATION NUMBER: 60/113, 296
PRIOR FILING DATE: December 22, 1998
PRIOR APPLICATION NUMBER: 60/146, 222
PRIOR FILING DATE: July 28, 1999
PRIOR APPLICATION NUMBER: PCT/US98/19330
PRIOR FILING DATE: September 16, 1998
PRIOR APPLICATION NUMBER: PCT/US98/25108
PRIOR FILING DATE: December 1, 1998

	U	Mismatches	0	Indels	0	Caps	0
QY	1	MRPLVLLLLGLAAGSPPLDDNKKIRSLCPGKPGGLPGKHGHSQGLPGCDGKGDGNDGARG	60				
Db	1	MRPLVLLLLGLAAGSPPLDDNKKIRSLCPGKPGGLPGKHGHSQGLPGCDGKGDGNDGARG	60				
QY	61	APGKGGGGRPLGPGRGDDPGRGAGAPGCTGPGAGCSCVPSRPSAFAKRSRSPVPSPD	120				
Db	61	APGKGGGGRPLGPGRGDDPGRGAGAPGCTGPGAGCSCVPSRPSAFAKRSRSPVPSPD	120				
QY	121	APLPGRVLVYNNQGHYDAVTGKFTQVPGVGYRYVAVHATYRYRASLDQFDLVKKGESIASFFQ	180				
Db	121	APLPGRVLVYNNQGHYDAVTGKFTQVPGVGYRYVAVHATYRYRASLDQFDLVKKGESIASFFQ	180				
QY	181	FFGCGMPKPAASLSGCAAMVRLPEDEQVWVQVGYDYLGIYASIKTOSTESGFLVYSDMHSSP	240				
Db	181	FFGCGMPKPAASLSGCAAMVRLPEDEQVWVQVGYDYLGIYASIKTOSTESGFLVYSDMHSSP	240				
QY	241	VFA 243					
Db	241	VFA 243					

RESULT 4
US-09-944-944-42
Sequence 42, Application US/099444944
Patent No. US20020173463A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David

1 APPLICANT: Eaton, Dan
 2 APPLICANT: Ferrara, Napoleone
 3 APPLICANT: Filvaroff, Ellen
 4 APPLICANT: Gerritsen, Mary
 5 APPLICANT: Goddard, Audrey
 6 APPLICANT: Godowski, Paul
 7 APPLICANT: Grimaldi, Christopher
 8 APPLICANT: Gurney, Austin
 9 APPLICANT: Hillan, Kenneth
 10 APPLICANT: Kljavin, Ivar
 11 APPLICANT: Napier, Mary
 12 APPLICANT: Roy, Margaret
 13 APPLICANT: Tumas, Daniel
 14 APPLICANT: Wood, William
 15 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 16 FILE REFERENCE: P2348P1
 17 CURRENT FILING DATE: 2001-09-26
 18 CURRENT FILING DATE: 2001-09-26
 19 PRIOR APPLICATION NUMBER: 09/866,028
 20 PRIOR FILING DATE: 2001-05-25
 21 PRIOR APPLICATION NUMBER: 60/067,411
 22 PRIOR FILING DATE: December 3, 1997
 23 PRIOR APPLICATION NUMBER: 60/069,334
 24 PRIOR FILING DATE: December 11, 1997
 25 PRIOR APPLICATION NUMBER: 60/069,335
 26 PRIOR FILING DATE: December 11, 1997
 27 PRIOR APPLICATION NUMBER: 60/069,278
 28 PRIOR FILING DATE: December 11, 1997
 29 PRIOR APPLICATION NUMBER: 60/069,425
 30 PRIOR FILING DATE: December 12, 1997
 31 PRIOR APPLICATION NUMBER: 60/069,656
 32 PRIOR FILING DATE: December 16, 1997
 33 PRIOR APPLICATION NUMBER: 60/069,654
 34 PRIOR FILING DATE: December 16, 1997
 35 PRIOR APPLICATION NUMBER: 60/069,702
 36 PRIOR FILING DATE: December 16, 1997
 37 PRIOR APPLICATION NUMBER: 60/069,870
 38 PRIOR FILING DATE: December 17, 1997
 39 PRIOR APPLICATION NUMBER: 60/069,873
 40 PRIOR FILING DATE: December 17, 1997
 41 PRIOR APPLICATION NUMBER: 60/068,017
 42 PRIOR FILING DATE: December 18, 1997
 43 PRIOR APPLICATION NUMBER: 60/070,440
 44 PRIOR FILING DATE: January 5, 1998
 45 PRIOR APPLICATION NUMBER: 60/074,086
 46 PRIOR FILING DATE: February 9, 1998
 47 PRIOR APPLICATION NUMBER: 60/074,092
 48 PRIOR FILING DATE: February 9, 1998
 49 PRIOR APPLICATION NUMBER: 60/075,945
 50 PRIOR FILING DATE: February 25, 1998
 51 PRIOR APPLICATION NUMBER: 60/112,850
 52 PRIOR FILING DATE: December 16, 1998
 53 PRIOR APPLICATION NUMBER: 60/113,296
 54 PRIOR FILING DATE: December 22, 1998
 55 PRIOR APPLICATION NUMBER: 60/146,222
 56 PRIOR FILING DATE: July 28, 1999
 57 PRIOR APPLICATION NUMBER: PCT/US98/19330
 58 PRIOR FILING DATE: September 16, 1998
 59 PRIOR APPLICATION NUMBER: PCT/US98/25108
 60 PRIOR FILING DATE: December 1, 1998
 61 PRIOR APPLICATION NUMBER: 09/216,021
 62 PRIOR FILING DATE: December 16, 1998
 63 PRIOR APPLICATION NUMBER: 09/218,517
 64 PRIOR FILING DATE: December 22, 1998
 65 PRIOR APPLICATION NUMBER: 09/254,311
 66 PRIOR FILING DATE: March 3, 1999
 67 PRIOR APPLICATION NUMBER: PCT/US99/12252
 68 PRIOR FILING DATE: June 22, 1999
 69 PRIOR APPLICATION NUMBER: PCT/US99/21090
 70 PRIOR FILING DATE: September 15, 1999
 71 PRIOR APPLICATION NUMBER: PCT/US99/28409
 72 PRIOR FILING DATE: No. US20020173463a1bchoc 70 1000

```
;; PRIOR APPLICATION NUMBER: PCT/US99/28313
;; PRIOR FILING DATE: No. US20020173463A:ember 30, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/28301
;; PRIOR FILING DATE: December, 1999
;; PRIOR APPLICATION NUMBER: PCT/US99/30095
;; PRIOR FILING DATE: December 16, 1999
;; PRIOR APPLICATION NUMBER: PCT/US00/03565
;; PRIOR FILING DATE: February 11, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/04414
;; PRIOR FILING DATE: February 22, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/05841
;; PRIOR FILING DATE: March 2, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/08439
;; PRIOR FILING DATE: March 30, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/14042
;; PRIOR FILING DATE: May 22, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/20710
;; PRIOR FILING DATE: July 28, 2000
;; PRIOR APPLICATION NUMBER: PCT/US00/32678
;; PRIOR FILING DATE: December 1, 2000
;; PRIOR APPLICATION NUMBER: PCT/US01/06520
;; NUMBER OF SEQ ID NOS: 120
;; SEQ ID NO 42
;; LENGTH: 243
;; TYPE: PRT
;; ORGANISM: Homo Saplen
US-09-944-944-42
```

```
Query Match      100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPTGPHHGSOGLPGRDGRDGRDAPG 60
DB      1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPTGPHHGSOGLPGRDGRDGRDAPG 60
QY      61 APGEKGEGRGRLPGPRDPPRGAGPAGPTGPGAGCSVPFRSAFSAKRSSESRVPPSD 120
DB      61 APGEKGEGRGRLPGPRDPPRGAGPAGPTGPGAGCSVPFRSAFSAKRSSESRVPPSD 120
QY      121 APLPFDRLVNEQGHYDAVTGKFTCOVPGVYFVAVHATYRASLOFDLVKNGESIASFQ 180
DB      121 APLPFDRLVNEQGHYDAVTGKFTCOVPGVYFVAVHATYRASLOFDLVKNGESIASFQ 180
QY      181 FFGGMPKASLSGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSFGFLVYSDWHSSP 240
DB      181 FFGGMPKASLSGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSFGFLVYSDWHSSP 240
QY      241 VFA 243
DB      241 VFA 243
```

RESULT 5

```
US-09-944-907-42
; Sequence 42, Application US/09944907
; Publication No. US20020198147A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Baton, Dan
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Hillan, Kenneth
; APPLICANT: Kijavlin, Ivar
; APPLICANT: Napier, Mary
; APPLICANT: Roy, Margaret
```

```
;; APPLICANT: Tumas, Daniel
;; APPLICANT: Wood, William
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE OF INVENTION: ACIDS ENCODING THE SAME
;; FILE REFERENCE: P2548P1C1
;; CURRENT APPLICATION NUMBER: US/09/944,907
;; PRIOR APPLICATION NUMBER: 2001-08-31
;; PRIOR FILING DATE: 2001-05-25
;; NUMBER OF SEQ ID NOS: 120
;; SEQ ID NO 42
;; LENGTH: 243
;; TYPE: PRT
;; ORGANISM: Homo Saplen
US-09-944-907-42
```

```
Query Match      100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPTGPHHGSOGLPGRDGRDGRDAPG 60
DB      1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPGHPGLPTGPHHGSOGLPGRDGRDGRDAPG 60
QY      61 APGEKGEGRGRLPGPRDPPRGAGPAGPTGPGAGCSVPFRSAFSAKRSSESRVPPSD 120
DB      61 APGEKGEGRGRLPGPRDPPRGAGPAGPTGPGAGCSVPFRSAFSAKRSSESRVPPSD 120
QY      121 APLPFDRLVNEQGHYDAVTGKFTCOVPGVYFVAVHATYRASLOFDLVKNGESIASFQ 180
DB      121 APLPFDRLVNEQGHYDAVTGKFTCOVPGVYFVAVHATYRASLOFDLVKNGESIASFQ 180
QY      181 FFGGMPKASLSGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSFGFLVYSDWHSSP 240
DB      181 FFGGMPKASLSGAMVRLPEPDQVWVQVGVGYIGIYASIKTDSFGFLVYSDWHSSP 240
QY      241 VFA 243
DB      241 VFA 243
```

RESULT 6

```
US-09-944-929-42
; Sequence 42, Application US/09944929
; Publication No. US20020197612A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin
; APPLICANT: Botstein, David
; APPLICANT: Baton, Dan
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul
; APPLICANT: Grimaldi, Christopher
; APPLICANT: Gurney, Austin
; APPLICANT: Hillan, Kenneth
; APPLICANT: Kijavlin, Ivar
; APPLICANT: Napier, Mary
; APPLICANT: Roy, Margaret
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P2548P1C1
; CURRENT APPLICATION NUMBER: US/09/944,929
; PRIOR APPLICATION NUMBER: 2001-08-31
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 120
; SEQ ID NO 42
; LENGTH: 243
; TYPE: PRT
```

ORGANISM: Homo Sapien
US-09-944-929-42

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7,9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLGLAGSPPDDNKIPSLCPGHPGLPTGPHGHSOGRLGRDGRGRGAG 60
DB 1 MRPLVLLGLAGSPPDDNKIPSLCPGHPGLPTGPHGHSOGRLGRDGRGRGAG 60
QY 61 APGKGGGRRGPGPRGDPGPRGEGAPGPTGPGAGECSVPKRSARSRVPPSD 120
DB 61 APGKGGGRRGPGPRGDPGPRGEGAPGPTGPGAGECSVPKRSARSRVPPSD 120
QY 121 APPLFDRVLNBOGHDVATGKFTCPVGYFAVATYRASLOPDLKNESTASFFQ 180
DB 121 APPLFDRVLNBOGHDVATGKFTCPVGYFAVATYRASLOPDLKNESTASFFQ 180
QY 181 FFGGMRKPSLSGCAVRLPEPDQVWVGVGDYIGITVASIKTSTFGFLYSDMHSP 240
DB 181 FFGGMRKPSLSGCAVRLPEPDQVWVGVGDYIGITVASIKTSTFGFLYSDMHSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 7

US-10-028-072-362
Sequence 362, Application US/10028072
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Mei-Qiang
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang
TITLE OR INVENTION:
FILE REFERENCE:
CURRENT APPLICATION NUMBER: US/10/028,072
CURRENT FILING DATE: 2001-12-19
PRIOR APPLICATION NUMBER: 60/049911
PRIOR FILING DATE: 1997-06-18
PRIOR APPLICATION NUMBER: 60/056974
PRIOR FILING DATE: 1997-08-26
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059115
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059117
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059122
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059184
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/059263
PRIOR FILING DATE: 1997-09-18
PRIOR APPLICATION NUMBER: 60/059352
PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059588

PRIOR FILING DATE: 1997-09-19
PRIOR APPLICATION NUMBER: 60/059836
PRIOR FILING DATE: 1997-09-24
PRIOR APPLICATION NUMBER: 60/062250
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062285
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/062814
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/062816
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063045
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063082
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/063127
PRIOR FILING DATE: 1997-10-24
PRIOR APPLICATION NUMBER: 60/063327
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063329
PRIOR FILING DATE: 1997-10-27
PRIOR APPLICATION NUMBER: 60/063550
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063561
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/063704
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063733
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063735
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063738
PRIOR FILING DATE: 1997-10-29
PRIOR APPLICATION NUMBER: 60/063755
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/064248
PRIOR FILING DATE: 1997-11-03
PRIOR APPLICATION NUMBER: 60/064809
PRIOR FILING DATE: 1997-11-07
PRIOR APPLICATION NUMBER: 60/065186
PRIOR FILING DATE: 1997-11-12
PRIOR APPLICATION NUMBER: 60/065846
PRIOR FILING DATE: 1997-11-17
PRIOR APPLICATION NUMBER: 60/066364
PRIOR FILING DATE: 1997-11-21
PRIOR APPLICATION NUMBER: 60/066453
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066511
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/066770
PRIOR FILING DATE: 1997-11-24
PRIOR APPLICATION NUMBER: 60/069212
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069278
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069334
PRIOR FILING DATE: 1997-12-11
PRIOR APPLICATION NUMBER: 60/069694
PRIOR FILING DATE: 1997-12-16
PRIOR APPLICATION NUMBER: 60/072320
PRIOR FILING DATE: 1998-01-23
PRIOR APPLICATION NUMBER: 60/073612
PRIOR FILING DATE: 1998-02-04
PRIOR APPLICATION NUMBER: 60/074086
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/074092
PRIOR FILING DATE: 1998-02-09
PRIOR APPLICATION NUMBER: 60/077791
PRIOR FILING DATE: 1998-03-12
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20

PRIOR APPLICATION NUMBER: 60/079294
 PRIOR FILING DATE: 1998-03-25
 PRIOR APPLICATION NUMBER: 60/079663
 PRIOR FILING DATE: 1998-02-27
 PRIOR APPLICATION NUMBER: 60/079728
 PRIOR FILING DATE: 1998-03-27
 PRIOR APPLICATION NUMBER: 60/080165
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/081203
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081229
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081695
 PRIOR FILING DATE: 1998-04-14
 PRIOR APPLICATION NUMBER: 60/081817
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081818
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082999
 PRIOR FILING DATE: 1998-04-24
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083545
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085149
 PRIOR FILING DATE: 1998-05-12
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085697
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085704
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/086414
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/086430
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088730
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088741
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089332
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445

PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090538
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/091360
 PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07

Query Match 100.0%; Score 243; DB 9; Length 243;
 Best Local Similarity 100.0%; Pred. No. 7.9e-202;
 Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLILGLAAGSPPLDDNKIPSLCGHGLPETPGHSGQLPGRDGHDGARG 60
 DB 1 MRPLVLLILGLAAGSPPLDDNKIPSLCGHGLPETPGHSGQLPGRDGHDGARG 60
 QY 61 APGEKEGGRPGIPGRGDPGRGEGAPGAPGEGSVPPRSAPSAKRSRVPSPSD 120
 DB 61 APGEKEGGRPGIPGRGDPGRGEGAPGAPGEGSVPPRSAPSAKRSRVPSPSD 120
 QY 121 APPEPRVLVNEQGHDAVTGKFTQVPCVYFAVAHATYRASLQFDLVKNSEIASFFQ 180
 DB 121 APPEPRVLVNEQGHDAVTGKFTQVPCVYFAVAHATYRASLQFDLVKNSEIASFFQ 180
 QY 181 FFGWMPKPSLSGAMVRLPEPDYVWVGVCDYIGIYASTIDTSGFLYVSMHSSP 240
 DB 181 FFGWMPKPSLSGAMVRLPEPDYVWVGVCDYIGIYASTIDTSGFLYVSMHSSP 240
 QY 241 VFA 243
 DB 241 VFA 243

RESULT 8

US-10-121-049-362
 Sequence 362, Application US/10121049
 Publication No. US20030022239A1

GENERAL INFORMATION:

APPLICANT: Beresini, Kevin P.
 APPLICANT: Deforge, Laura
 APPLICANT: Desnoyers, Luc
 APPLICANT: Filvaroff, Ellen
 APPLICANT: Gao, Wei-Qiang
 APPLICANT: Gerritsen, Mary E.
 APPLICANT: Goddard, Audrey
 APPLICANT: Godowski, Paul J.
 APPLICANT: Gurney, Austin L.
 APPLICANT: Sherwood, Steven
 APPLICANT: Smith, Victoria
 APPLICANT: Stewart, Timothy A.
 APPLICANT: Tumas, Daniel
 APPLICANT: Watanabe, Colin K
 APPLICANT: Wood, William
 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 FILE REFERENCE: P33081C17
 CURRENT APPLICATION NUMBER: US/10/121,049
 PRIOR APPLICATION removed - See File Wrapper or Palm
 NUMBER OF SEQ ID NOS: 550
 SEQ ID NO 362
 LENGTH: 243
 TYPE: PRT
 ORGANISM: Homo Sapien
 US-10-121-049-362

Query Match 100.0%; Score 243; DB 9; Length 243;

Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPHGLPCTGCHHSOGLPGRDGRDGDAGP 60
DB 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPHGLPCTGCHHSOGLPGRDGRDGDAGP 60
QY 61 APGEKGGRRPGLPGRGDPGRGAGPAGTGPAGECVPPRASFSAKRSESRVPPSPD 120
DB 61 APGEKGGRRPGLPGRGDPGRGAGPAGTGPAGECVPPRASFSAKRSESRVPPSPD 120
QY 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVNKGESIASFFQ 180
DB 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVNKGESIASFFQ 180
QY 181 FFGGMPKPRASISGAMVRLPEDDVWVQVGVGYIGIYASIKTDSFSGFLVYSMDHSSP 240
DB 181 FFGGMPKPRASISGAMVRLPEDDVWVQVGVGYIGIYASIKTDSFSGFLVYSMDHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 9

US-10-123-904-362
Sequence 362, Application US/10123904
Publication No. US20030022328A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C54
CURRENT FILING DATE: 2002-04-16
PRIOR APPLICATION REMOVED - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-123-904-362

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPHGLPCTGCHHSOGLPGRDGRDGDAGP 60
DB 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPHGLPCTGCHHSOGLPGRDGRDGDAGP 60
QY 61 APGEKGGRRPGLPGRGDPGRGAGPAGTGPAGECVPPRASFSAKRSESRVPPSPD 120
DB 61 APGEKGGRRPGLPGRGDPGRGAGPAGTGPAGECVPPRASFSAKRSESRVPPSPD 120
QY 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVNKGESIASFFQ 180
DB 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVNKGESIASFFQ 180

DB 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVNKGESIASFFQ 180

QY 181 FFGGMPKPRASISGAMVRLPEDDVWVQVGVGYIGIYASIKTDSFSGFLVYSMDHSSP 240
DB 181 FFGGMPKPRASISGAMVRLPEDDVWVQVGVGYIGIYASIKTDSFSGFLVYSMDHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 10

US-10-140-470-362
Sequence 362, Application US/10140470
Publication No. US2003002231A1
GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: Deforge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerltzen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C160
CURRENT FILING DATE: 2002-05-06
PRIOR APPLICATION REMOVED - See File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-470-362

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPHGLPCTGCHHSOGLPGRDGRDGDAGP 60
DB 1 MRPLVLLLLGLAAGSPPLDDNKIPSLCPHGLPCTGCHHSOGLPGRDGRDGDAGP 60
QY 61 APGEKGGRRPGLPGRGDPGRGAGPAGTGPAGECVPPRASFSAKRSESRVPPSPD 120
DB 61 APGEKGGRRPGLPGRGDPGRGAGPAGTGPAGECVPPRASFSAKRSESRVPPSPD 120
QY 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVNKGESIASFFQ 180
DB 121 APLEPDRVLVNEQGHYDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVNKGESIASFFQ 180
QY 181 FFGGMPKPRASISGAMVRLPEDDVWVQVGVGYIGIYASIKTDSFSGFLVYSMDHSSP 240
DB 181 FFGGMPKPRASISGAMVRLPEDDVWVQVGVGYIGIYASIKTDSFSGFLVYSMDHSSP 240
QY 241 VFA 243
DB 241 VFA 243

RESULT 11
US-09-796-753-68

```

; Sequence 68, Application US/09796753
; Publication No. US20030027998A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
; FILE REFERENCE: 7853-227-999
; CURRENT FILING DATE: US/09/796,753
; PRIOR FILING DATE: 2001-03-01
; PRIOR APPLICATION NUMBER: 09/183,175
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 09/223,094
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/223,546
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/224,246
; PRIOR FILING DATE: 1998-12-30
; PRIOR APPLICATION NUMBER: 09/259,388
; PRIOR FILING DATE: 1999-02-26
; PRIOR APPLICATION NUMBER: 60/122,458
; PRIOR FILING DATE: 1999-03-01
; PRIOR APPLICATION NUMBER: 09/312,359
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/336,536
; PRIOR FILING DATE: 1999-06-18
; PRIOR APPLICATION NUMBER: 09/342,687
; PRIOR FILING DATE: 1999-06-29
; PRIOR APPLICATION NUMBER: 09/345,464
; PRIOR FILING DATE: 1999-06-30
; PRIOR APPLICATION NUMBER: 09/365,164
; PRIOR FILING DATE: 1999-07-30
; PRIOR APPLICATION NUMBER: 09/399,723
; PRIOR FILING DATE: 1999-09-20
; PRIOR APPLICATION NUMBER: 09/409,634
; PRIOR FILING DATE: 1999-09-30
; PRIOR APPLICATION NUMBER: 09/471,179
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 09/474,071
; PRIOR FILING DATE: 1999-12-29
; PRIOR APPLICATION NUMBER: 09/514,010
; PRIOR FILING DATE: 2000-02-25
; PRIOR APPLICATION NUMBER: 09/516,745
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: 09/572,002
; PRIOR FILING DATE: 2000-05-14
; PRIOR APPLICATION NUMBER: 09/597,993
; PRIOR FILING DATE: 2000-06-19
; PRIOR APPLICATION NUMBER: 09/599,596
; PRIOR FILING DATE: 2000-06-22
; PRIOR APPLICATION NUMBER: 09/630,334
; PRIOR FILING DATE: 2000-07-31
; PRIOR APPLICATION NUMBER: 09/606,565
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/606,317
; PRIOR FILING DATE: 2000-06-29
; PRIOR APPLICATION NUMBER: 09/665,666
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: 09/677,751
; PRIOR FILING DATE: 2000-09-30
; NUMBER OF SEQ ID NOS: 162
; SEQ ID NO 68
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-796-753-68

```

```

Query Match      100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MRPLVLLGLAAGSPPLDNNKIPSLCPGHRLPCTPGHSGSLPGRDGRDGAAG 60
|||||

```

```

Db      1 MRPLVLLGLAAGSPPLDNNKIPSLCPGHRLPCTPGHSGSLPGRDGRDGAAG 60
Qy      61 APGEKGGRRGRLPGRRGDPGPRGEGAPGPTGAGECVPPRSASFSAKRSESRVPPSD 120
Db      61 APGEKGGRRGRLPGRRGDPGPRGEGAPGPTGAGECVPPRSASFSAKRSESRVPPSD 120
Qy      121 APLEFDRVLVNEQGHDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESTIASFFQ 180
Db      121 APLEFDRVLVNEQGHDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESTIASFFQ 180
Qy      181 FRGGMKPRASLSGAMVRLPEPDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
Db      181 FRGGMKPRASLSGAMVRLPEPDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
Qy      241 VPA 243
Db      241 VPA 243

```

```

RESULT 12
US-10-175-746-362
; Sequence 362, Application US/10175746
; Publication No. US20030027270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C353
; CURRENT APPLICATION NUMBER: US/10/175,746
; PRIOR FILING DATE: 2002-06-19
; PRIOR APPLICATION REMOVED - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
; US-10-175-746-362

```

```

Query Match      100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MRPLVLLGLAAGSPPLDNNKIPSLCPGHRLPCTPGHSGSLPGRDGRDGAAG 60
|||||
1 MRPLVLLGLAAGSPPLDNNKIPSLCPGHRLPCTPGHSGSLPGRDGRDGAAG 60
61 APGEKGGRRGRLPGRRGDPGPRGEGAPGPTGAGECVPPRSASFSAKRSESRVPPSD 120
61 APGEKGGRRGRLPGRRGDPGPRGEGAPGPTGAGECVPPRSASFSAKRSESRVPPSD 120
121 APLEFDRVLVNEQGHDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESTIASFFQ 180
121 APLEFDRVLVNEQGHDAVTGKFTCOVPGVYFAVHATVYRASLQFDLVKNGESTIASFFQ 180
181 FRGGMKPRASLSGAMVRLPEPDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240
181 FRGGMKPRASLSGAMVRLPEPDQVWVGVDYIGIYASIKTDSFGFLVYSDMHSSP 240

```

QY 241 VFA 243
Db 241 VFA 243

RESULT 13

US-10-176-918-362
; Sequence 362, Application US/10176918
; Publication No: US20030027275A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C382
; CURRENT APPLICATION NUMBER: US/10/176,918
; PRIOR APPLICATION: removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-176-918-362

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHGSQGLPGRDGRGRCGAPG 60
Db 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHGSQGLPGRDGRGRCGAPG 60
QY 61 APGEKGGGRGRLGPRGDDPRGEGAGPAGPTGAGECSVPPRSASAKRSSESRVPPSPD 120
Db 61 APGEKGGGRGRLGPRGDDPRGEGAGPAGPTGAGECSVPPRSASAKRSSESRVPPSPD 120
QY 121 APLEPDRVLNVEGCHYDAVTGKFTQVPGVYFAVHATYRASLQFDLYKNESIASFEQ 180
Db 121 APLEPDRVLNVEGCHYDAVTGKFTQVPGVYFAVHATYRASLQFDLYKNESIASFEQ 180
QY 181 FFGGMPKRPASLSGGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFSGFLYSDWHSSP 240
Db 181 FFGGMPKRPASLSGGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFSGFLYSDWHSSP 240
QY 241 VFA 243
Db 241 VFA 243

RESULT 14

US-10-176-921-362
; Sequence 362, Application US/10176921
; Publication No: US20030027276A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen

APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C288
CURRENT APPLICATION NUMBER: US/10/176,921
PRIOR APPLICATION: removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 362
LENGTH: 243
TYPE: PRT
ORGANISM: Homo Sapien
US-10-176-921-362

Query Match 100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Caps 0;

QY 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHGSQGLPGRDGRGRCGAPG 60
Db 1 MRPLVLLLLGLAAGSPPLDNDKIPSLCPGHPGLPTGPHHGSQGLPGRDGRGRCGAPG 60
QY 61 APGEKGGGRGRLGPRGDDPRGEGAGPAGPTGAGECSVPPRSASAKRSSESRVPPSPD 120
Db 61 APGEKGGGRGRLGPRGDDPRGEGAGPAGPTGAGECSVPPRSASAKRSSESRVPPSPD 120
QY 121 APLEPDRVLNVEGCHYDAVTGKFTQVPGVYFAVHATYRASLQFDLYKNESIASFEQ 180
Db 121 APLEPDRVLNVEGCHYDAVTGKFTQVPGVYFAVHATYRASLQFDLYKNESIASFEQ 180
QY 181 FFGGMPKRPASLSGGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFSGFLYSDWHSSP 240
Db 181 FFGGMPKRPASLSGGAMVRLPEPDQVWVGVGDYIGIYASIKTDSFSGFLYSDWHSSP 240
QY 241 VFA 243
Db 241 VFA 243

RESULT 15

US-10-137-865-362
; Sequence 362, Application US/10137865
; Publication No: US20030032155A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K

```

; APPLICANT: Wood, William
; APPLICANT: Zhang, Zhenlin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C154
; CURRENT APPLICATION NUMBER: US/10/137,865
; CURRENT FILING DATE: 2002-05-03
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 362
; LENGTH: 243
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-137-865-362

```

```

Query Match      100.0%; Score 243; DB 9; Length 243;
Best Local Similarity 100.0%; Pred. No. 7.9e-202;
Matches 243; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MRPLVLLLLGLAAGSPPLDDNKTIPSLCPGHPGLPCTPGHSGSGILPGRDGRDGAAG 60
   |||
Db 1 MRPLVLLLLGLAAGSPPLDDNKTIPSLCPGHPGLPCTPGHSGSGILPGRDGRDGAAG 60
   |||
QY 61 APGEKGGGRRPGILPGRGDPGPRGEGAGPAGPTGPAEGCSVPPRSASFSAKRSESRVPPSD 120
   |||
Db 61 APGEKGGGRRPGILPGRGDPGPRGEGAGPAGPTGPAEGCSVPPRSASFSAKRSESRVPPSD 120
   |||
QY 121 APLPFDRLVNEQGHYDAVTGKFTCOVPGVYTFEAVHATVYRASLOFDLVKNGESTIASFFQ 180
   |||
Db 121 APLPFDRLVNEQGHYDAVTGKFTCOVPGVYTFEAVHATVYRASLOFDLVKNGESTIASFFQ 180
   |||
QY 181 FFGGMPKRPASLSGCAVRLPEPDQVWVGVGYIYASIKTDSTFSGFLVYSDWHSSP 240
   |||
Db 181 FFGGMPKRPASLSGCAVRLPEPDQVWVGVGYIYASIKTDSTFSGFLVYSDWHSSP 240
   |||
QY 241 VFA 243
   |||
Db 241 VFA 243

```

Search completed: June 18, 2003, 15:15:19
 Job time : 54 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2003, 02:43:56 ; Search time 3598 Seconds

(without alignments)
11138.011 Million cell updates/sec

Title: US-09-943-851A-41

Perfect score: 1377

Sequence: 1 gactagctctctctgagctc.....aaaaaaaaaaaaaaaaaaaa 1377

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenBank:
1: gb-ba:*
2: gb-hcg:*
3: gb-in:*
4: gb-om:*
5: gb-ov:*
6: gb-pat:*
7: gb-ph:*
8: gb-pl:*
9: gb-pr:*
10: gb-ro:*
11: gb-sts:*
12: gb-sy:*
13: gb-un:*
14: gb-vl:*
15: em-ba:*
16: em-fun:*
17: em-hum:*
18: em-in:*
19: em-mu:*
20: em-om:*
21: em-or:*
22: em-ov:*
23: em-pat:*
24: em-ph:*
25: em-pl:*
26: em-ro:*
27: em-sts:*
28: em-un:*
29: em-vl:*
30: em-hcg-hum:*
31: em-hcg-in:*
32: em-hcg-liv:*
33: em-hcg-mus:*
34: em-hcg-pin:*
35: em-hcg-rod:*
36: em-hcg-mam:*
37: em-hcg-vrt:*
38: em-sy:*
39: em-hcg-hum:*
40: em-hcg-mus:*
41: em-hcg-other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1377	100.0	1377	6	AX464228
2	1347.4	97.9	1370	9	BC029485
3	1322.2	96.0	1337	9	AF329841
4	1322.2	96.0	1347	6	AR138193
5	1064.4	77.3	1068	9	HSW800923
6	909.8	66.1	191362	9	AP003396
7	909.8	66.1	219574	9	AP003396
8	896.2	65.1	182429	2	AP001557
9	896.2	50.6	4220	10	AF469650
10	694	50.4	173038	2	AP001003
11	693.6	50.4	1234	10	BC025174
12	691.2	50.2	1271	10	BC023068
13	612.8	44.5	198902	2	AP001156
14	513.4	37.3	173657	2	AC112557
15	504	36.6	729	6	AR138195
16	500.4	36.3	249283	2	AC124577
17	472.4	34.3	150653	2	AC107174
18	439.4	31.9	441	9	HUM292F04
19	187.4	13.6	150653	2	AC107174
20	161.8	11.8	2288	9	AB055132
21	107	7.8	1086	5	AB067770
22	100.6	7.3	4443	9	AK074129
23	100.6	7.3	4908	6	AX430953
24	100.6	7.3	128133	9	AL138787
25	100.6	7.3	173817	2	AC012141
26	99.6	7.2	855	6	AX039874
27	97.4	7.1	1938	9	HUM286GLE
28	92.6	6.7	152147	9	AC091842
29	92.6	6.7	153248	9	AC112191
30	91.6	6.7	1161	6	AX039965
31	91.6	6.7	1176	9	AF329836
32	91.6	6.7	2404	9	BC011699
33	91.4	6.6	1134	4	AF269230
34	90.6	6.6	947	10	BC028770
35	90.6	6.6	1276	6	AR034252
36	90.6	6.6	1276	6	AX195209
37	90.6	6.6	1276	6	AX358519
38	90.6	6.6	1276	6	AX358519
39	90.2	6.6	194674	2	AC122477
40	90	6.5	132884	2	AC128403
41	88.2	6.4	11354	2	AC129431
42	88	6.4	1357	9	AX068378
43	88	6.4	1393	9	AF329838
44	88	6.4	159766	9	AC104942
45	86.8	6.3	178850	9	AL359736

ALIGNMENTS

RESULT 1
AX464228
LOCUS
DEFINITION
Sequence 361 from Patent W00140466.
VERSION
AX464228.1 GI:21899130
KEYWORDS
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
Baker,K.P., Beresini,M., DeForge,L., Desnoyers,L., Filvaroff,E.,
Gao,W.O., Gerritsen,M.E., Goddard,A., Godowski,P.J., Gurney,A.L.,
Sherwood,S., Smith,V., Stewart,T.A., Tunas,D., Watanabe,C.K.,

667 **ГЛАВНОЕ УПРАВЛЕНИЕ ЗАЩИТЫ ПРАВОСудИЯ**

```

FEATURES
source      .
            Eastlake Ave. East, Seattle, WA 98102, USA
            Location/Qualifiers
            1..1337
            /organism="Homo sapiens"
            /db_xref="taxon:9606"

```


40 CGCAGGAGCCGAACGAGACTGGGGTGAACGGCCAGGGGGGCGCTTGCCGGGAGA 99
 11 CGAGAGGGAGCGAACCAGAGACTGGGGTGACGGCAGGGGGGCGCTTGCCGGGAGA 70
 100 AGGCGGGGGCTTGGACACCACTGAGGGGTCGAGAGAGAGCCGCCCTCCGGGTGC 159
 71 AGCGGGGGGCTGGAGACCACTGGAGGGTCCGGAGTAGGAGAGGCCCGAAGAG 130
 160 GCCATGGGGAGCCGGAGGGGGGACTGAGAGAGACCCGGGCTCCGGGTCCGGGTGC 219
 131 GCCATGGGGAGCCGGAGGGGGAGCTCGAGAGAGACCCGGGCTCCGGGTGC 190
 220 CAGCGGTATGAGGCCACTCTCTGCTCTGCTCTCTGGGCTGGCGGCGCGCTCGGCCCC 279
 191 CAGCGGTATGAGGCCACTCTCTCTCTCTCTCTCTGGGCGTGGCGGCGCGCTCGGCCCC 250
 280 ACTGAGACGACAAGATGCCAGCTCTGCGCGGGGACCCCGGCTTCCAGCAGCC 339
 251 ACTGAGACGACAAGATGCCAGCTCTGCGCGGGGACCCCGGCTTCCAGCAGCC 310
 340 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCGCGAGTGGCGGAGGGCGGAGCGGCG 399
 311 GGGCCACCATGGCAGCCAGGGCTTGCCGGGCGCGAGTGGCGGAGGGCGGAGCGGCG 370
 400 GCCCGGGGCTCCGGGAGAGAAAGGCGGAGGCGGAGCGCGGACTCCGGACTCGAGG 459
 371 GCCCGGGGCTCCGGGAGAGAAAGGCGGAGGCGGAGCGGAGCGGAGCTCCGGACTCGAGG 430
 460 GGACCCCGGGCGGAGAGAGCGGGACCGCGGGGGCCACCGGGGCTCGCGGGGAGTG 519
 431 GGACCCCGGGCGGAGAGAGCGGGACCGCGGGGGCCACCGGGGCTCGCGGGGAGTG 490
 520 CTGGGTCCTCCGGATCCGCTTCAAGCGCCAAAGCCCTCGAGAGCCGGAGTCCCTCCGCC 579
 491 CTGGGTCCTCCGGATCCGCTTCAAGCGCCAAAGCCCTCGAGAGCCGGAGTCCCTCCGCC 550
 580 GTCTGAGCAGCCCTTGGCGCTTGACCGGCGCTGGTGGTGAACGAGCGAGGAGATTACAGCG 639
 551 GTCTGAGCAGCCCTTGGCGCTTGACCGGCGCTGGTGGTGAACGAGCGAGGAGATTACAGCG 610
 640 CGTCAACCGGCAAGTTCACCTGCGAGGTGGTGGTCTCTACTCTTCCGCGCATATGCGAC 699
 611 CGTCAACCGGCAAGTTCACCTGCGAGGTGGTGGTCTCTACTCTTCCGCGCATATGCGAC 670
 700 CGTCAACCGGCAAGTTCACCTGCGAGGTGGTGGTCTCTACTCTTCCGCGCATATGCGCTTT 759
 671 CGTCAACCGGCAAGTTCACCTGCGAGGTGGTGGTCTCTACTCTTCCGCGCATATGCGCTTT 730
 760 CTTCAAGTTTTTCGGGGGGTGCGCAACCGACGCTCTCGGGGGGGCGCATGTGAG 819

11 CGAGAGGAGCGAACCCAGGACTGGGTGACGGCAGGCGAGGGGCCCTGGCCGGGAGA 70

Db 11 CGAGAGGAGGACACACAGGACTGTTGGGTGACGGCAGGCGAGGGGGGCGCTTGGCCGGGGGAGA 70

QY 100 AGCGCGGGGGCTGGAGCACACCAACTGGAGGCTCCGAGTAGCGAGGAGGCGCCCCCAAGAGAC 150

D	b	71	AGCGCGGGGGCTGAGACACCACCACTGAGGGGCTCGGAGTAGACGACGCCCCCAAGAG	130
Q	y	160	GCCATCGGGAGCGCGGAGGGGAGACTGGCAAGAGACCCCGCGTCCGGGCTCCGGTGC	219
D	b	131	GCCATCGGGAGCGCGGAGGGGAGACTGGCAAGAGACCCCGCGTCCGGGCTCCGGTGC	190
Q	y	220	CAGCGGTATGAGAGCCACTCTCTGCTCGTCTGCTCTGAGGCTGGCGCGCGTCCGCC	279
D	b	191	CAGCGGTATGAGAGCCACTCTCTGCTCTGCTCTGAGGCTGGCGCGCGTCTGCC	250
Q	y	280	ACTGAGCAACAAGATCCCGAGGCTCTGGCCGGGGACCCGGGCTTCAGGACGCC	339
D	b	251	ACTGAGCAACAAGATCCCGAGGCTCTGGCCGGGGACCCGGGCTTCAGGACGCC	310
Q	y	340	GGGACCAATGGCAACCGAGGCTTCGGGGGCGGATGGCGGAGCGCGGACGACGCC	399
D	b	311	GGGACCAATGGCAACCGAGGCTTCGGGGGCGGATGGCGGAGCGCGGACGACGCC	370
Q	y	400	GCCCGGGCTCCGGGAGAGAAAGCGAGGCGGGAGCGGGACCTCCGGGACCTGAG	459
D	b	371	GCCCGGGCTCCGGGAGAGAAAGCGAGGCGGGAGCGGGACCTCCGGGACCTGAG	430
Q	y	460	GGACCCCGGCGCGGAGAGAGCGGGAGACCCCGGGGCCACCGGGCTCCGGGAGTG	519
D	b	431	GGACCCCGGCGCGGAGAGAGCGGGAGACCCCGGGGCCACCGGGCTCCGGGAGTG	490
Q	y	520	CTCGGTGCTCCGGGATCCGCTTCAGAGCGCAAGCGCTCCGAGAGCGGGTCCCGCC	579
D	b	491	CTCGGTGCTCCGGGATCCGCTTCAGAGCGCAAGCGCTCCGAGAGCGGGTCCCGCC	550
Q	y	580	GTCGAGAGCACCCCTTGCCCTTCGACCGCGCTCGTGGTGAACGAGCGGACATTACGAGC	639
D	b	551	GTCGAGAGCACCCCTTGCCCTTCGACCGCGCTCGTGGTGAACGAGCGGACATTACGAGC	610
Q	y	640	CGTCACCGGGAAGTTACCTCCAGGTGCGTGGGTCTACTACTTCGCGCGCATATGGCAC	699
D	b	611	CGTCACCGGGAAGTTACCTCCAGGTGCGTGGGTCTACTACTTCGCGCGCATATGGCAC	670
Q	y	700	GCTCTACCGGGCAGCGCTGACTGTTGATCTGGTGAAGATGGCAATGCCATGCTCTT	759
D	b	671	GCTCTACCGGGCAGCGCTGACTGTTGATCTGGTGAAGATGGCAATGCCATGCTCTT	720
Q	y	760	CTTCCAGTTTTCGGGGGTGGGCCCAAGCAGCCCTCGCTCGGGGGGGCCATATGGTAG	819
D	b	731	CTTCCAGTTTTCGGGGGTGGGCCCAAGCAGCCCTCGCTCGGGGGGGCCATATGGTAG	790
Q	y	820	GCTGAGCGCTGAGGACCAAGTGGGTGACGTGGGTGGGTGATCAATTGGCATCTTA	879
D	b	791	GCTGAGCGCTGAGGACCAAGTGGGTGACGTGGGTGGGTGATCAATTGGCATCTTA	850
Q	y	880	TGCGGCAATCAAGACAGACAGACACTTCTCGGATTTCTTGTTACTTCCGACTGGGACAG	939
D	b	851	TGCGGCAATCAAGACAGACAGACACTTCTCGGATTTCTTGTTACTTCCGACTGGGACAG	910
Q	y	940	CTCCCAAGCTTTTCTAGTGCACAGCAAGGAGCTCATGCTCTCACTCTTGAAG	999
D	b	911	CTCCCAAGCTTTTCTAGTGCACAGCAAGGAGCTCATGCTCTCACTCTTGAAG	970
Q	y	1000	AGGCTGAGGAGCTTACAAACAGGTCAATCAGGAGGGCTGGCCCTGGAAATATTGTAA	1059
D	b	971	AGGCTGAGGAGCTTACAAACAGGTCAATCAGGAGGGCTGGCCCTGGAAATATTGTAA	1030
Q	y	1060	TGACTAGGAGGTGGGTAGAGCACTCCGTTCTCTCTCTGCGCAAGAAATGGAAAGT	1119
D	b	1031	TGACTAGGAGGTGGGTAGAGCACTCCGTTCTCTCTCTGCGCAAGAAATGGAAAGT	1090
Q	y	1120	GGGCTGTGCGAATCAGGTCTGGGAGATGGGCAAGTGGCTGATTTCTGCCAAGACAG	1179
D	b	1091	GGGCTGTGCGAATCAGGTCTGGGAGATGGGCAAGTGGCTGATTTCTGCCAAGACAG	1150
Q	y	1180	AGGAGTGTGCTGTGTCGCAAGTGAATGATCCCGCATGTTGCTGGTCCAGAGCCACGG	1239

Accession	Sequence	Position
Db	1151 AGAGGTGTGCTGTGCTGGCAAGTGTAAAGTCCCGCAGTTGCTGTGTCAGAGAGCCACGG	12110
QY	1240 TGGGGTCTCTCTTTCCTGATGCTCTGCTGTGATCCGCCACCCCTCTGTCTCT	12995
Db	1211 TGGGGTGTCTCTCTCTGCTCTCTGATCTCTCCACCCCTCTCTCTCT	12707
QY	1300 GGGGCGCGCCCTTTTCTCAGAGTACTCATTAACCTAAGAACCTCATTAACCAAAAAA	13595
Db	1271 GGGGCGCGCCCTTTTCTCAGAGTACTCATTAACCTAAGAACCTCAAAAAA	13300
QY	1360 AAAAAA 1366	
Db	1331 AAAAAA 1337	

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	FEATURES
RESULT 5 HSM600923							
LOCUS	HSM600923	1068 bp	mRNA	linear			
DEFINITION	Homo sapiens mRNA; cDNA DKFZ588B0621 (from clone DKFZ588B0621).						
ACCESSION	AL110261						
VERSION	AL110261.1						
KEYWORDS	GI:5817222						
SOURCE	Homo sapiens.						
ORGANISM							

REFERENCE
AUTHORS

TITLE	JOURNAL	COMMENT
Wiemann, S. Direct Submission Submitted (15-AUG-1999) MIPS, Am Klopferplatz 18a, D-82152 Martinsried, GERMANY		
Clone from S. Wiemann, Molecular Genome Analysis Center		

FEATURES

Research Center (DKFZ): Email: s.wienemann@dkfz-heidelberg.de;
sequencing consortium of the German Genome Project. This clone
(DKFZ568B0621) is available at the RZPD in Berlin. Please contact
the RZPD: Ressourcenzentrum@heubnerweg.6.14059
Berlin-Charlottenburg, GERMANY; Email: clonerezpd.de Further
information about the clone and the sequencing project is available
at <http://www.mips.biochem.mpg.de/proj/cDNA>.
Location@onalfiers

SOURCE

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKFZp586B0621"
/tissue_type="uterus"
/clone_id="586 (synonym: hntel). Vector pSport1; host
DH10B; sites NotI + SalI/MluI"
/dev_stage="adult"
1..660
/gene="DKFZp586B0621"
<1..660
CDS

```

```

/gene="DKFZP586B0621"
/note="similarity to complement subcomponent C1q chain C"
/codon_start=1
/product="hypothetical protein"
/protein_id="CAB53702.1"
/db_xref="GI:3817223"
/translation="PSLCPGHRLPGTGRHHGSGQLPGRCGRGRDGRDAGAPAGKGECE
GRGLGPRGPDGVRGGEAPGAPGPTGAGCESTVPPRAAFSGKSSVPSADPLPRT
RLVLDNGHGVDVAVTGRKFCYQPGVGYVPAVAVTAITLFDILKNGKSEVPSADPLPRT
WKPASLISGGAWVRLEPEDQWVQVGYGVDTIGITLIVASIKRDTSFSGFLVYSIMSHSPF
A"

```

polya_signal	1031.	1036	
polya_site	1049		
BASE COUNT	187 a	334 c	349 g
ORIGIN			198 t

Query Match	77.3%	Score 1064.4	DB 9	Length 1068
Best Local Similarity	99.9%	Pred. No. 9.2e-174		
Matches 1065; Conservative	0	Mismatches 1	Indels 0	Gaps 0

OY 299 CCGAGCTCTGCCCCGGGCGACCCCGGCTTCCAGGACGCGCGCCACCATTGACGACGAG 358
 DB 1 CCGAGCTCTGCCCCGGGCGACCCCGGCTTCCAGGACGCGCGCCACCATTGACGACGAG 60
 OY 359 GCGTTCCGCGGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 418
 DB 61 GCGTTCCGCGGCGCGCATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
 OY 419 AAGGCGAGGCGGCGGAGCGCGGAGCTGCGGGAGCTCGAGGAGGAGCGCGCGCGCGGAG 478
 DB 121 AAGGCGAGGCGGCGGAGCGCGGAGCTGCGGGAGCTCGAGGAGGAGCGCGCGCGCGGAG 180
 OY 479 GAGGCGGAGCGCGGCGGCG 538
 DB 181 GAGGCGGAGCGCGGCGGCG 240
 OY 539 GCGTTCCAGCG 598
 DB 241 GCGTTCCAGCG 300
 OY 599 TTGCGACCG 658
 DB 301 TTGCGACCG 360
 OY 659 TGCCAGGTCCTGGGCGTCTACTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 718
 DB 361 TGCCAGGTCCTGGGCGTCTACTACTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420
 OY 719 CAGTTTGAATGCTGTAAGAAATGCGGAAATCCATTGCTCTTCTTCTTCTTCTTCTTCT 778
 DB 421 CAGTTTGAATGCTGTAAGAAATGCGGAAATCCATTGCTCTTCTTCTTCTTCTTCTTCT 480
 OY 779 TGCCCGAAGCG 838
 DB 481 TGCCCGAAGCG 540
 OY 839 GTGTGGTGCAGGTGCGTGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 898
 DB 541 GTGTGGTGCAGGTGCGTGTGGTGAATGATGATGATGATGATGATGATGATGATGATGAT 600
 OY 899 AGCAGCTTCTCCGAGATTTCTGCTGCTACTCCGACTGCGACAGCTCCCGACTCTTCTCT 958
 DB 601 AGCAGCTTCTCCGAGATTTCTGCTGCTACTCCGACTGCGACAGCTCCCGACTCTTCTCT 660
 OY 959 TGCCCACTGCAAGTAGTGCATGCTCTACTCTCTGAAGAGAGGCTGTAGGCTGACAAAC 1018
 DB 661 TGCCCACTGCAAGTAGTGCATGCTCTACTCTCTGAAGAGAGGCTGTAGGCTGACAAAC 720
 OY 1019 CAGGTCATCCAGAGAGGCGTGGCGCCCTCGAATATTTGTAATGACTAGAGAGAGGCTGA 1078
 DB 721 CAGGTCATCCAGAGAGGCGTGGCGCCCTCGAATATTTGTAATGACTAGAGAGAGGCTGA 780
 OY 1079 GAGCACTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1138
 DB 781 GAGCACTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 OY 1139 TGCGACATGAGGCGAGTGGCTGATTTCTGCGCAAGACAGAGAGAGTGTGCTGCTGCTG 1198
 DB 841 TGCGACATGAGGCGAGTGGCTGATTTCTGCGCAAGACAGAGAGAGTGTGCTGCTGCTG 900
 OY 1199 AAGTGAATGCTCCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1258
 DB 901 AAGTGAATGCTCCCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
 OY 1259 TCCTGTGCTTCTGATGCTCTCCAGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1318
 DB 961 TCCTGTGCTTCTGATGCTCTCCAGCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 OY 1319 GAGATCACTCAATAACTAAGAACCTCATTAATAAAAAAAAAAAAAA 1364
 DB 1021 GAGATCACTCAATAACTAAGAACCTCATTAATAAAAAAAAAAAAAA 1066

RESULT 6
 AP003396/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 HOMO SAPIENS DNA, clone:RP11-334E6.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 1
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens genomic DNA
 Published only in Database (2002)
 2 (bases 1 to 191362)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (12-MAR-2001) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 Tel:81-45-503-9111, Fax:81-45-503-9170
 (E-mail:hattori@gs.c.riken.go.jp, URL:http://hnp.gsc.riken.go.jp/
 This work was done in collaboration with Arai, Y., Kubo, T. and
 Ohki, M.
 National Cancer Center Research Institute
 Cancer Genomic Division
 5-1-1, Tsukiji, Chuo-ku, Tokyo, JAPAN
 zip: 104-0045
 phone: 81-3-3542-2511 ex 4752, fax: 81-3-3542-0688 e-mail:
 yarai@nci.go.jp.
 Location/Qualifiers
 source
 1..191362
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11g"
 /clone="RP11-334E6"
 BASE COUNT 46535 a 45977 c 48722 g 50128 t
 ORIGIN
 Query Match 66.1%; Score 909.8; DB 9; Length 191362;
 Best Local Similarity 99.8%; Pred. No. 2.6e-147;
 Matches 911; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 OY 438 CCGGACTGCGCGGACCTCGAGGGGACCCCGCGCGGAGAGAGCGGACCGCGGGCG 497
 DB 6384 CAGGACTGCGCGGACCTCGAGGGGACCCCGCGGAGAGAGCGGACCGCGGGCG 6325
 OY 498 CCACGGGCGCGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 557
 DB 6324 CCACGGGCGCGCGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 6265
 OY 558 CCGAGAGCCCGGATGCTCCGCGCTGAGCGACACCTTGGCTTGCAGCGCGGTGCTGTA 617
 DB 6264 CCGAGAGCCCGGATGCTCCGCGCTGAGCGACACCTTGGCTTGCAGCGCGGTGCTGTA 6205
 OY 618 ACAGAGCGGACATTTACGACCGCGTACCGCGCAAGTTCACTGCGAGTGCCTGGGCTCT 677
 DB 6204 ACAGAGCGGACATTTACGACCGCGTACCGCGCAAGTTCACTGCGAGTGCCTGGGCTCT 6145
 OY 678 ACTACTGCGCGCTCATATGACACCGCTTACCGGCGCGAGCTGAGTTGATGTGGGAGA 737
 DB 6144 ACTACTGCGCGCTCATATGACACCGCTTACCGGCGCGAGCTGAGTTGATGTGGGAGA 6085
 OY 738 ATGGCGAATCATATGCTCTTTCTTTCAGTTTTCGGGGGTGGCCCAAGCCAGCCTGCG 797
 DB 6084 ATGGCGAATCATATGCTCTTTCTTTCAGTTTTCGGGGGTGGCCCAAGCCAGCCTGCG 6025
 OY 798 TCTCGGGGGGCGCATGTGAGGCTGAGACCTGAGACCAAGTGTGGTGCAGGTGGCTG 857

```

Db      6024 TCTCGGGGGGGCCATGCTGAGAGCTGAGACCTGAGAGACCAAGTGTGGGTGAGGTGGTG 5965
QY      858 TGGGTGACTACTTGGCATCTATGCGACATCAGACAGACAGACCTTCTCCGATTTC 917
Db      5964 TGGGTGACTACTTGGCATCTATGCGACATCAGACAGACAGACCTTCTCCGATTTC 5905
QY      918 TGGGTGACTACTTGGCATCTATGCGACATCAGACAGACAGACCTTCTCCGATTTC 977
Db      5904 TGGGTGACTACTTGGCATCTATGCGACATCAGACAGACAGACCTTCTCCGATTTC 5845
QY      978 TCAATGCTCTCACTCTAGAGAGAGGTGTGAGAGCTGAGACAGAGTATGAGAGAGGCT 1037
Db      5844 TCAATGCTCTCACTCTAGAGAGAGGTGTGAGAGCTGAGACAGAGTATGAGAGAGGCT 5785
QY      1038 GGGCCCCCTGCAATATTTGTGAATGACTAGAGAGGTGGGTGAGAGACACTCTCTCTCT 1097
Db      5784 GGGCCCCCTGCAATATTTGTGAATGACTAGAGAGGTGGGTGAGAGACACTCTCTCTCT 5725
QY      1098 GCTGGCAAGAAATGGAGAGTGGCTGTCTGCGATCAGAGTCTGGCAGCATGGGAGTGG 1157
Db      5724 GCTGGCAAGAAATGGAGAGTGGCTGTCTGCGATCAGAGTCTGGCAGCATGGGAGTGG 5665
QY      1158 CTGGATTCTTGGCCAGACAGAGAGTGTGCTGTGCTGCGCAAGTAAAGTCCCGAGTT 1217
Db      5664 CTGGATTCTTGGCCAGACAGAGAGTGTGCTGTGCTGCGCAAGTAAAGTCCCGAGTT 5605
QY      1218 GCTGTGCTCCAGAGAGCCAGAGTGGGTGCTCTCTCTCTGCTGCTGCTCTCTGATC 1277
Db      5604 GCTGTGCTCCAGAGAGCCAGAGTGGGTGCTCTCTCTCTGCTGCTGCTCTCTGATC 5545
QY      1278 CTCCCAACCCCTCTCTCTCTCTGCGGCGGCGCTTCTGAGAGATACCAATAAACT 1337
Db      5544 CTCCCAACCCCTCTCTCTCTCTGCGGCGGCGCTTCTGAGAGATACCAATAAACT 5485
QY      1338 AAGAACCTCTATA 1350
Db      5484 AAGAACCTCTATA 5472

RESULT 7
AP002956/c 219574 bp DNA linear PRI 15-AUG-2002
LOCUS      Homo sapiens genomic DNA, chromosome 11q clone:R1105h09, complete
sequence.
ACCESSION  AP002956
VERSION     AP002956.1 GI:22255355
KEYWORDS    HGC.
SOURCE      Homo sapiens DNA, clone:R1105h09.
ORGANISM    Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1
AUTHORS     Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE       Homo sapiens 219,574 genomic DNA of 11q
JOURNAL     Published Only in Database (2002)
REFERENCE   2 (bases 1 to 219574)
AUTHORS     Hattori,M., Toyoda,A., Taylor,T.D., Fujiyama,A., Yada,T.,
            Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE       Direct Submision
JOURNAL     Submitted (24-NOV-2000) Masahira Hattori, The Institute of Physical
            and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
            1-7-22, Saitiro-chou,Tsuri-ku, Yokohama, Kanagawa 230-0045, Japan
            (E-mail:hattori@gs.c.riken.go.jp, URL:http://hgp.gs.c.riken.go.jp/,
            Tel:81-45-503-9111, Fax:81-45-503-9170)
            Overlapping Clones: PD71A4, PQ38G8
            Cytogenetic Position: 21q22.1, region:D21S226-AML
            STS Markers (ePCR): sts633747, SHGC-16045, D11S2450, SHGC-7143,
            WI-16056,
            sts631054, sts-X57110, WI-8978, GDB:198117, sts629284, SHGC-130657,
            D11S1941E, WI-14589, sts660191, sts650535, A002048, sts626946,
            sts641092,

```

```

TIGR-A005D38, bdaa7902, D11S869E, RH80030
Contamination: none detected
non-ACGT bases: none
Additional author information
Arai,Y., Kudo,T., Onki,M.
National Cancer Center Research Institute
Cancer Genomic Division
5-1-1, Tsukiji, Chuo-Ku, Tokyo, JAPAN
zip: 104-0045
phone: +81-3-3542-2511 ex4752, fax: +81-3-3542-0688 e-mail:
yara@ncc.go.jp.
FEATURES
Source      Location/Qualifiers
            1..219574
            /organism="Homo sapiens"
            /db_xref="taxon:9606"
            /chromosome="11"
            /map="11q"
            /clone="R1105h09"
BASE COUNT  54883 a 50356 c 53844 g 60491 t
ORIGIN
Query Match      66.1%; Score 909.8; DB 9; Length 219574;
Best Local Similarity 99.8%; Pred. No.2,6e-147;
Matches 911; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY      438 CGGAGTCCCGGAGCTCGAGAGGAGCCCGGCGCGGAGAGAGGCGGAGCCGCGGCG 497
Db      209122 CAGGACTGCGGGAGCTCGAGAGGAGCCCGGCGCGGAGAGAGGAGGAGCCGCGGCG 209063
QY      498 CCACCGGCGCTCGGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 557
Db      209062 CCACCGGCGCTCGGGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 209003
QY      558 CCGAGAGCGGGGTGCTCGCGCGCTGAGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 617
Db      209002 CCGAGAGCGGGGTGCTCGCGCGCTGAGAGCAAGCTGCTGCTGCTGCTGCTGCTGCTG 208943
QY      618 ACAGAGAGGAGATTAAGAGCGCGCTGACCGGAGTTCACTGCGAGTGTGCTGCGTCT 677
Db      208942 ACAGAGAGGAGATTAAGAGCGCGCTGACCGGAGTTCACTGCGAGTGTGCTGCGTCT 208883
QY      678 ACTACTGCGCGCTGATGCGAGCGCTGACCGGAGTTCACTGCGAGTGTGCTGCGTCTGCTG 737
Db      208882 ACTACTGCGCGCTGATGCGAGCGCTGACCGGAGTTCACTGCGAGTGTGCTGCGTCTG 208823
QY      738 ATGCGGAATCAATTCCTCTCTTCTTCAAGTCTTCTGCGGGGGTGGCCAAAGCAGCTGCG 797
Db      208822 ATGCGGAATCAATTCCTCTCTTCTTCAAGTCTTCTGCGGGGGTGGCCAAAGCAGCTGCG 208763
QY      798 TCTCGGGGGGGCCATGCTGAGAGCTGAGAGCTGAGAGCAAGTGTGGGTGAGGTGGTG 857
Db      208762 TCTCGGGGGGGCCATGCTGAGAGCTGAGAGCTGAGAGCAAGTGTGGGTGAGGTGGTG 208703
QY      858 TGGGTGACTACTTGGCATCTATGCGACATCAGACAGACAGACCTTCTCCGATTTC 917
Db      208702 TGGGTGACTACTTGGCATCTATGCGACATCAGACAGACAGACCTTCTCCGATTTC 208643
QY      918 TGGGTGACTACTTGGCATCTATGCGACATCAGACAGACAGACCTTCTCCGATTTC 977
Db      208642 TGGGTGACTACTTGGCATCTATGCGACATCAGACAGACAGACCTTCTCCGATTTC 208583
QY      978 TCAATGCTCTCACTCTAGAGAGAGGTGTGAGAGCTGAGACAGAGTCAATCCAGAGAGGCT 1037
Db      208582 TCAATGCTCTCACTCTAGAGAGAGGTGTGAGAGCTGAGACAGAGTCAATCCAGAGAGGCT 208523
QY      1038 GGGCCCCCTGCAATATTTGTGAATGACTAGAGAGGTGGGTGAGAGACACTCTCTCTCTCT 1097
Db      208522 GGGCCCCCTGCAATATTTGTGAATGACTAGAGAGGTGGGTGAGAGACACTCTCTCTCTCTCT 208463
QY      1098 GCTGGCAAGAAATGGAGAGTGGCTGTGCTGCGATCAGAGTCTGGCAGCATGGGAGTGG 1157
Db      208462 GCTGGCAAGAAATGGAGAGTGGCTGTGCTGCGATCAGAGTCTGGCAGCATGGGAGTGG 208403

```

QY 1158 CTGATTTCTGCCAGACAGAGAGTCTGCTGTCAGTAAATGCCCAAGTT 1217
 Db 208402 CTGATTTCTGCCAGACAGAGAGTCTGCTGTCAGTAAATGCCCAAGTT 208343
 QY 1218 GCTGTGTCAGAGAGCCAGCGTGGGTCCTCTCTCTGCTGCTGCTGATC 1277
 Db 208342 GCTGTGTCAGAGAGCCAGCGTGGGTCCTCTCTCTGCTGCTGCTGATC 208283
 QY 1278 CTCCACACCCCTCTCTCTCTGCGGCGCCCTTTCTCAGAGATCAGTCAATTAACCT 1337
 Db 208282 CTCCACACCCCTCTCTCTCTGCGGCGCCCTTTCTCAGAGATCAGTCAATTAACCT 208223
 QY 1338 AAGAACCTCATTA 1350
 Db 208222 AAGAACCTCATTA 208210

RESULT 8
 LOCUS AP001557 182429 bp DNA linear HTG 03-NOV-2000
 DEFINITION Homo sapiens chromosome 11 clone RP11-680A7 map 11q23, WORKING
 ACCESSION AP001557 3 GI:11094164
 VERSION HTG; HTG_PHASE1; HTG_DRAFT.
 KEYWORDS Homo sapiens DNA, clone:RP11-680A7.
 SOURCE Homo sapiens
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1 (bases 1 to 182429)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens 182,429 genomic DNA of 11q23
 2 (bases 1 to 182429)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P., Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Submitted (29-MAR-2000) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170) On Nov 3, 2000 this sequence version replaced gi:18117391.
 COMMENT
 Center: RIKEN Genomic Sciences Center(GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 Project Information
 Center project name: HumDRAFT11
 Center clone name: RP11-680A7
 Summary Statistics
 Sequencing Vector: PCR products; 100% of reads
 Chemistry: Dye-terminator; Eppendorf; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 175597 bases at least Q40
 Consensus quality: 178555 bases at least Q30
 Consensus quality: 179768 bases at least Q20
 Insert size: 180429; sum-of-ctrls
 Quality coverage: 9.01x in Q20 bases; sum-of-ctrls

NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

1 30107 contig of 30107 bp in length
 30208 54357 contig of 24150 bp in length
 54458 69839 contig of 15382 bp in length
 69940 84048 contig of 14109 bp in length

FEATURES

source

84149 95177 contig of 11029 bp in length
 95278 107763 contig of 12466 bp in length
 107864 119057 contig of 11194 bp in length
 119158 128134 contig of 8977 bp in length
 128235 136903 contig of 8669 bp in length
 137004 145887 contig of 8884 bp in length
 145988 153199 contig of 7212 bp in length
 153300 159565 contig of 6266 bp in length
 159666 163580 contig of 3915 bp in length
 163581 165152 contig of 1472 bp in length
 165153 165252 contig of 100 bp in length
 165253 166901 contig of 1649 bp in length
 166902 170460 contig of 3459 bp in length
 170461 172761 contig of 2201 bp in length
 172762 172861 contig of 2744 bp in length
 172862 175605 contig of 2603 bp in length
 175606 175705 contig of 100 bp in length
 175706 178308 contig of 2603 bp in length
 178309 178408 contig of 100 bp in length
 178409 180918 contig of 2510 bp in length
 180919 181018 contig of 100 bp in length
 181019 182429 contig of 1411 bp in length.

NOTE: This is a 'working draft' sequence. It currently consists of 21 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of 'N', but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1. 182429
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="11"
 /map="11q23"
 /clone="RP11-680A7"

[illegible]

Query Match	Score	DB 2;	Length
Best Local Similarity	65.1%;	896.2;	182429;
	99.6%;	Pred No 5	8e-145;

	Y	Nucleotides	3'	Indels	1'	Gaps	1'
OY	438	CGGACATCCGGGGACTTCGAGGGAGCCCCGGGGCCGGAGAGAGCGGGAGACCCTCGGGGC					497
Db	108011	CAGCACTGCCGGGACCTTCGAGGGAGACCCCGGGCCCGGAGAGAGCGGGACCCCGGGGC					108070
OY	498	CCACCGGGCCTGC CGGGAGATGCTCGTGCCTCCGGAATCCGGCTTCAGCCCAAAGCCT					557
Db	108071	CCACCGGGCCTGC CGGGAGATGTCTCGTCTCTCCGGATTCCGCTTCAGCCCAAAGCCT					108130
OY	558	CCGAGACCGGGGTCCTCCCGCTCTGACGCAACCCCTTGCCCTTGACCGCGTGGTGTGA					617
Db	108131	CCGAGACCGGGGTCCTCCCGCTCTGACGCAACCCCTTGCCCTTGACCGCGTGGTGTGA					108190
OY	618	ACGAGCAGGAGACATTTACGACGCGCTCACCGGGAATTAACCTGCAGAGTCCCTGGGCTCT					677
Db	108191	ACGAGCA - GGACATTATCGACGCGCGTACACGGCAAGTTCACTGCAAGTCCCGGGGCTCT					108249
OY	678	ACTACTTGCCTCCCATGCGAACCGTCTACCGGGCCAGCGTCGAGTTTGATCTGTGTAAGA					737
Db	108250	ACTACTTGCCTCCCATGCGAACCGTCTACCGGGCCAGCGTTTGATCTGTGTAAGA					108309
OY	738	ATGGCGAATCATTTGCCTCTTTCTTCAGTTTTTCGGGGGTGGCCCAAGCCAAGCCTCGC					797
Db	108310	ATGGCGAATCATTTGCCTCTTTCTTCAGTTTTTCGGGGGTGGCCCAAGCCAAGCCTCGC					108369

QY	798	TCCTCGGGGGGGCCCATGGTGAAGGCTGGAGCGCTGAGGACCAACTGTGGGTGACAGTGGGTG	857
Db	108370	TCCTCGGGGGGGGGCCCATGGTGAAGGCTGGAGCGCTGAGGACCAACTGTGGGTGACAGTGGGTG	108423
QY	858	TGGGTGACTACATTGGGCACTCATATGGCCAGCATCAAGACAGACAGCAACCTTCTCCGGATTTC	917
Db	108430	TGGGTGACTACATTGGGCACTCATATGGCCAGCATCAAGACAGACAGCAACCTTCTCCGGATTTC	108488
QY	918	TGGTGTACTCCGACTGGCAGACAGCTCCGCCAGTCTTTTCTTAGTGGCCCACTGCAAAGTAGC	977
Db	108490	TGGTGTACTCCGACTGGCAGACAGCTCCGCCAGTCTTTTCTTAGTGGCCCACTGCAAAGTAGC	108549
QY	978	TCATGCTCTCACTCTTGAAAGAGGGGTGTGAGGCTACACAAACAGGTCTATCAGAGAGGCT	1037
Db	108550	TCATGCTCTCACTCTTGAAAGAGGGGTGTGAGGCTACACAAACAGGTCTATCAGAGAGGCT	108609
QY	1038	GGCCCCCTCGAATATTTGTAATGATATAGGAGGTGGGGTGAAGACACTCTCCGTCGTCT	1097
Db	108610	GGCCCCCTCGAATATTTGTAATGATATAGGAGGTGGGGTGAAGACACTCTCCGTCGTCT	108669
QY	1098	GCTGGCAAGGAATGGGACAGTGGCTGTGTGGATCAGAGCTGTGGCAGCATGGGGCAAGTG	1157
Db	108670	GCTGGCAAGGAATGGGACAGTGGCTGTGTGGATCAGAGCTGTGGCAGCATGGGGCAAGTG	108729
QY	1158	CTGATATTTCTGCCCAAGACACAGAGAGTGTGCTGTGCTGGCAAGTGTAAATCCCCAGTT	1217
Db	108730	CTGATATTTCTGCCCAAGACACAGAGAGTGTGCTGTGCTGGCAAGTGTAAATCCCCAGTT	108789
QY	1218	GCCTGTGCTCCAGAGAGCCACAGGTGGGGGTGCTCTCTCTGTCCTCTGCTTCTCTGATC	1277
Db	108790	GCCTGTGCTCCAGAGAGCCACAGGTGGGGGTGCTCTCTCTCTGTCCTCTGCTTCTCTGATC	108849
QY	1278	CTCCCCACCCCTCTGCTCTGTGGGGCCGGCCCTTTTCTAGAGATCACTCAATAAACCT	1337
Db	108850	CTCCCCACCCCTCTGCTCTGTGGGGCCGGCCCTTTTCTAGAGATCACTCAATAAACCT	108909
QY	1338	AAGAACCCTCAATA 1350	
Db	108910	AAGAACCCTCAATA 108922	

```

RESULT 9
AF469650
LOCUS
DEFINITION
AF469650 4220 bp mRNA linear ROD 31-JUL-2002
COMMENT: c1q tumor necrosis factor-related protein (Mfrp) and mRNAs, complete cds.
ACCESSION
AF469650
VERSION
AF469650.1 GI:22023766
KEYWORDS
SOURCE
Mus musculus.
ORGANISM
Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 4220)
Kameya,S., Hawes,N.L., Chang,B., Heckenlively,J.R., Naggert,J.K. and Nishina,P.M.
Mfrp, a gene encoding a frizzled related protein, is mutated in the mouse retinal degeneration 6
Hum. Mol. Genet. 11 (16), 1879-1886 (2002)
JOURNAL
12140190
PUBMED
2 (bases 1 to 4220)
Kameya,S., Naggert,J.K. and Nishina,P.M.
Direct Submission
JOURNAL
Submitted (17-JAN-2002) The Jackson Laboratory, 600 Main Street, Bar Harbor, ME 04609, USA
FEATURES
location/Qualifiers
1..4220
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/chromosome="9"
/map="26 cm"
source

```

```

gene      52..1806
          /gene="Mfrp"
CDS       52..1806
          /gene="Mfrp"
          /note="MFRP"
          /codon_start=1
          /product="membrane-type frizzled-related protein"
          /protein_id="AA089216.1"
          /db_xref="GI:22023767"
          /translation="MKDYDYLIRPEASELSKTEFCNPADPDEAPSPAPALORVY
SRLOAPMHAQRLRGLOPCDHFSEMCILLISLLISLLGLAVIAYIAOLQATSLRYTK
NPILRLGTLPKGVIPSTPTNTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTTPTT
FSPNPYDPLVPSHCYWHITOVAAQOTIOLKIALSTESMLTCLFDLELISPTPL
LAVGKIDPRLPLNTNTSHLRVSEYSDNDVESGQAWYQAVAGHMSCANHEFDCL
LCRLRSDYCDITLCAQSDSDANCSAKTLGCGNLTLGVFSTPNYPOHYAPPOOCT
WYIEVPVAGYRIEFNFHNSLEQAQEKFEYVEYEAISLGFSLFGKAPPEPLMY
SSMHLQAVIKETDGISSGFLATYQAINTESSCGPWERCGSGGARDLOMOMDLKD
CANDSNDNCSHLSPOPDLCEPVOYEMCGLISTNTAPINIVGATOTETVDILNG
YKSLTSLPCYOTFORFLCGLILVPRCTSLGTLILPCCRVCQAAEDQCCSSIALIGTWP
FNCRLPVAASTLEACSP"
          3097..3828
          /gene="C1qlnf5"
          3097..3828
          /gene="C1qlnf5"
          /note="C1QTNF5"
          /codon_start=1
          /product="complement-c1q tumor necrosis factor-related
protein"
          /protein_id="AA089217.1"
          /db_xref="GI:22023768"
          /translation="MRPLALLLILGLVSGSPPLDNDKIPICRGORGLPCTGHHGSO
GLFGDRGDRDQCAPCAPGPKGKGRGRLGPRGEPGREGAPKATIPAECSVP
RSAPFSAKRSERVPADPRLPDRVLNBQHTDPTTKETQVPGVYFVHAHY
RASLODVLVNGOSIASPFOYREGMPKPSLSGAAVNRLEPDDVWVQVGDYIGIT
ASIKTSTSGFLVYSDMHSSPYFA"
BASE COUNT      873 a 1282 c 1147 g 918 t
ORIGIN
Query Match      50.6%; Score 696.2; DB 10; Length 4220;
Best Local Similarity 78.0%; Pred. No. 2.3e-110;
Matches 921; Conservative 0; Mismatches 238; Indels 22; Gaps 6;
189 GAGAGACCCCGGCTCCGGGCTCCGGTGCCAGCGCTATGAGGCACACTCTGCTCTGC 248
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3059 GAGAGTTCCTGAGCTGAGCTCCGCGGTGACACCAATGAGGCGACATCTTGGCCCTTC 3118
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
249 TCGTCTGAGCGCTGCGCGGCTGCGCCCACTGAGAGCAACAAGATCCCAAGCTCT 308
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3119 TCGTCTGAGCTGCTGCTGCTCTGCTGCTGAGAGCAACAAGATCCCAAGCTCT 3178
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
309 GCCCGGGGACCCCGCTTCCAGGACCGCGGCGACCATGACGCGAGGCTTGGCGG 368
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3179 GTCGCGGCGACCGCGCTTCCAGGACACAGTCACCATGCGAGCCAGGCTTGGCTG 3238
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
369 GCCGCGATGGCGGCGCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 428
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3239 GCCGTGACGCGCGTATGCGCGCGGAGTCTGCGGTACCCCGAGATCGGAGAAAGCGG 3298
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
429 GCGGAGCGCGGAGCTGCGAGGAGACCGCGGCGCGGCGGAGAGAGCGCGGAG 488
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3299 GCGGAGAGACCGGACTACTGCGCCACACTGGGAGACCGCGGCGCGGAGAGCGCGG 3358
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
489 CCGCGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 548
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3359 CGATGGGGGCTATCGCGGCTGCGGCGGAGTCTGCGGTACCCCGAGATCGACTTCA 3418
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
549 CCAAGCGCTCGGAGCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 608
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3419 CCAAGCGATCGGAGACCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 3478
QY      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
609 TCGTGTGACGACGAGGACATACGAGCGCTTACCGCGGAGATTACCTGCGAGGTGC 668
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3479 TCGTGTGACGACGAGGCGCATACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGG 3538

```

```

QY      669 CTGGGCTACTACTTCGCGCTCCATGCCACCTCTACCGGGCGACGCTGAGTTGATC 728
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3539 CTGGCGCTACTACTTTCGCTGCGACGACGTCTACCGGGCGACGCTTGCAGTTGATC 3598
QY      729 TGGTGAAGATGCGCAATCATTTGCTCTTTTCTTCCAGTTTTCGGGGGGTGGCCAAAGC 788
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3599 TTCTCAAAAAGGCGGCTCATTCGCTCTTCTTCTTCCAGTATTTTGGGGGTGGCCAAAGC 3658
QY      789 CACCTGCTCTCGGGGGGGCGGACGCTGAGGCTGAGGCTGAGGACCAAGCTGGGTGC 848
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3659 CACCTGCTCTCGAGGGGTGCGATGTAAGCTGAAACCTGAGGACCAAGCTGGGTGC 3718
QY      849 AGTGGGTGCTGCTGCTACTGATTCATTTGATGAGATGAGGACGACAGCAAGCTTCT 908
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3719 AGTGGGTGCTGCTGATTAATTCATTTGATGAGATGAGGACGACAGCAAGCTTCT 3778
QY      909 CCGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 968
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3779 CTGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3838
QY      969 AAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1028
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3839 ACCCGAGCTGCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3898
QY      1029 AGGAGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1088
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3899 CAGGAGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3957
QY      1089 GCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1148
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3958 AGTCCACTGCTGCGCAATGATGATGAGACA--GCTGCTGAGTGA-----AGACAGCGTG 4011
QY      1149 GGGCAGTGGCGGATTTTGTGCGGACAGAGAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1208
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4012 GAGCAGTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4070
QY      1209 CCGCCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1265
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4071 C-----CTGGCCAGAGACTCCAGAGTGGATGCTCAATTCCTGCTGCTGCTGCTGCTGCT 4119
QY      1266 CTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1325
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4120 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4179
QY      1326 CTCATTAACCTAAGACCTCATATAAAAAAAAAAAAAAAAAAAAAA 1366
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
4180 CTCATTAACCTAAGACCTCATATAAAAAAAAAAAAAAAAAAAAAA 4220

```

RESULT 10
 AP001003
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AP001003 173038 bp DNA linear HTG 30-MAY-2000
 Homo sapiens chromosome 11 clone RP11-680C5 map 11q23, WORKING
 DRAFT SEQUENCE, 28 unordered pieces.
 AP001003
 AP001003.2 GI:117674
 HTG: HTGS-PHASE1; HTGS_DRAFT.
 Homo sapiens DNA, clone: RP11-680C5.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 1 (bases 1 to 173038)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens 173,038 genomic DNA of 11q23
 Published Only in Database (2000)
 2 (bases 1 to 173038)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (05-JAN-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC);

COMMENT

Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail: hatorilegsc.riken.go.jp,
URL: http://hgp.gsc.riken.go.jp/
Fax: 81-42-778-9923)
On May 31, 2000 this sequence version replaced g1:6997762.

Center: RIKEN Genomic Sciences Center(GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hatorilegsc.riken.go.jp

Project information

Center project name: HumDrafit1

Center clone name: RPI1-680C5

Summary Statistics

Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 149676 bases at least Q40
Consensus quality: 160868 bases at least Q30
Consensus quality: 167161 bases at least Q20
Insert size: 170338; sum-of-contigs
Quality coverage: 4.33x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```

1      15558 contig of 15558 bp in length
15659      33689 contig of 18031 bp in length
33790      47432 contig of 13643 bp in length
47533      63392 contig of 15860 bp in length
63493      74235 contig of 10742 bp in length
74335      85357 contig of 11023 bp in length
85458      93597 contig of 8140 bp in length
93698      101983 contig of 8286 bp in length
102084      110022 contig of 7939 bp in length
110123      118117 contig of 7994 bp in length
118217      126464 contig of 8248 bp in length
126555      132137 contig of 5572 bp in length
132237      135290 contig of 3054 bp in length
135391      140229 contig of 4839 bp in length
140330      143896 contig of 3576 bp in length
143997      147572 contig of 3576 bp in length
147673      150446 contig of 2774 bp in length
150447      153715 contig of 3169 bp in length
153716      157273 contig of 3458 bp in length
157274      159817 contig of 2444 bp in length
159818      161743 contig of 1826 bp in length
161744      163684 contig of 1841 bp in length
163685      165216 contig of 1432 bp in length
165217      166571 contig of 100 bp in length
166572      166757 contig of 1341 bp in length
166758      168307 contig of 1550 bp in length
168308      168407 contig of 100 bp in length
168408      170030 contig of 1623 bp in length
170031      170130 contig of 100 bp in length
170131      171725 contig of 1595 bp in length
171726      171825 contig of 100 bp in length
171826      173038 contig of 1213 bp in length

```

Sequence updated (26-May-2000).

NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1      15558 contig of 15558 bp in length
15559      15658 contig of 100 bp in length
15659      33689 contig of 18031 bp in length
33690      33789 contig of 100 bp in length
33790      47432 contig of 13643 bp in length
47433      47532 contig of 100 bp in length

```

FEATURES

SOURCE

```

misc_feature      47533      63392: contig of 15860 bp in length
misc_feature      63393      63492: gap of 100 bp
misc_feature      63493      74235: contig of 10742 bp in length
misc_feature      74235      74335: gap of 100 bp
misc_feature      74335      85357: contig of 11023 bp in length
misc_feature      85358      85457: gap of 100 bp
misc_feature      85458      93597: contig of 8140 bp in length
misc_feature      93598      93697: gap of 100 bp
misc_feature      93698      101983: contig of 8286 bp in length
misc_feature      101984      102083: gap of 100 bp
misc_feature      102084      110022: contig of 7939 bp in length
misc_feature      110023      110122: gap of 100 bp
misc_feature      110123      118116: contig of 7994 bp in length
misc_feature      118117      118216: gap of 100 bp
misc_feature      118217      126464: contig of 8248 bp in length
misc_feature      126465      126564: gap of 100 bp
misc_feature      126565      132136: contig of 5572 bp in length
misc_feature      132137      132236: gap of 100 bp
misc_feature      132237      135290: contig of 3054 bp in length
misc_feature      135291      135390: gap of 100 bp
misc_feature      135391      140229: contig of 4839 bp in length
misc_feature      140230      140329: gap of 100 bp
misc_feature      140330      143896: contig of 3576 bp in length
misc_feature      143897      143996: gap of 100 bp
misc_feature      143997      147572: contig of 3576 bp in length
misc_feature      147573      147672: gap of 100 bp
misc_feature      147673      150446: contig of 2774 bp in length
misc_feature      150447      150546: gap of 100 bp
misc_feature      150547      153715: contig of 3169 bp in length
misc_feature      153716      153815: gap of 100 bp
misc_feature      153816      157273: contig of 3458 bp in length
misc_feature      157274      157373: gap of 100 bp
misc_feature      157374      159817: contig of 2444 bp in length
misc_feature      159818      159917: gap of 100 bp
misc_feature      159918      161743: contig of 1826 bp in length
misc_feature      161744      161843: gap of 100 bp
misc_feature      161844      163684: contig of 1841 bp in length
misc_feature      163685      163784: gap of 100 bp
misc_feature      163785      165216: contig of 1432 bp in length
misc_feature      165217      165316: gap of 100 bp
misc_feature      165317      166571: contig of 1341 bp in length
misc_feature      166572      166757: gap of 100 bp
misc_feature      166758      168307: contig of 1550 bp in length
misc_feature      168308      168407: gap of 100 bp
misc_feature      168408      170030: contig of 1623 bp in length
misc_feature      170031      170130: gap of 100 bp
misc_feature      170131      171725: contig of 1595 bp in length
misc_feature      171726      171825: gap of 100 bp
misc_feature      171826      173038: contig of 1213 bp in length.

Location/Qualifiers
1. 173038
   /organism="Homo sapiens"
   /db_xref="taxon:9606"
   /chromosome="11"
   /map="11q23"
   /clone="RPI1-680C5"
1. 15558
   /note="assembly-fragment"
15659. 33689
   /note="assembly-fragment"
33790. 47432
   /note="assembly-fragment"
47533. 63392
   /note="assembly-fragment"
63493. 74234
   /note="assembly-fragment"
74335. 85357
   /note="assembly-fragment"
85458. 93597
   /note="assembly-fragment"
93698. 101983
   /note="assembly-fragment"
102084. 110022

```



```

misc_feature      /note="assembly_fragment"
110123..118116
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
118217..126464
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
126565..132136
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
132237..135290
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
135391..140229
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
140330..143896
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
143997..147572
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
147673..150446
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
150547..153715
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
153816..157273
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
157374..159817
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
159918..161743
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
161844..163684
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
163785..165216
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
165317..166657
/note="assembly_fragment"
misc_feature      /note="assembly_fragment"
166758..168307
/note="assembly_fragment"

```

Query Match 50.4%: Score 694; DB 2; Length 173038; Best Local Similarity 98.8%; Pred. No. 4e-110; Matches 741; Conservative 0; Mismatches 5; Indels 4; Gaps 4;

```

OY 604 CCGCGTCGTGTGAAGCAGCAGGAGCAATTACGAGCCCTCACCGCAAGTTCACCTGCCA 663
DB 166758 CCGCGTCGTGTGAAGCAGCAGGAGCAATTACGAGCCCTCACCGCAAGTTCACCTGCCA 166816
OY 664 GGTGCTGCGGTCTACTACTTCCGCGTCCATGCGACCGCTTACCGGGCCGACCGTCACTT 723
DB 166817 GGTGCTGCGGTCTACTACTTCCGCGTCCATGCGACCGCTTACCGGGCCGACCGTCACTT 166876
OY 724 TGATCTGTGAAGAAATGGCCAAATTCATTCCTTTTCCAGTTTTC-GGGGGGTGGC 782
DB 166877 TGATCTGTGAAGAAATGGCCAAATTCATTCCTTTTCCAGTTTTCGGGGGGGTGGC 166936
OY 783 CCAAGCCAGCCTCGCTTCGGGGGGGGCCATGCGAGCGCTGAGGCTGAGACCAAGTGT 842
DB 166937 CCAAGCCAGCCTCGCTTCGGGGGGGGCCATGCGAGCGCTGAGGCTGAGACCAAGTGT 166996
OY 843 GGTGCGAGGTGGTGTGGTGTGACTACATTCGATTCATGCGACCATCAAGACAGACAGCA 902
DB 166997 GGTGCGAGGTGGTGTGGTGTGACTACATTCGATTCATGCGACCATCAAGACAGACAGCA 167056
OY 903 CCTTCTCCGATTTCTGTGTACTCGGACGAGCCTCCCAAGCTTTTCTTACTAGTGC 962
DB 167057 CCTTCTCCGATTTCTGTGTACTCGGACGAGCCTCCCAAGCTTTTCTTACTAGTGC 167116
OY 963 CACTGCAAAAGTGAATGCTCTCACTCTCACTGAAGAGAGGAGGTGTGAGGCTGACCAACGAG 1022
DB 167117 CACTGCAAAAGTGAATGCTCTCACTCTCACTGAAGAGAGGAGGTGTGAGGCTGACCAACGAG 167176
OY 1023 TCAT-CCAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1080
DB 167177 TCAT-CCAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 167236
OY 1081 GCACCTTCCTGCTGCTGTGCGCAGGAATGGAACAGGCTGTGCTGATCAGATCTG 1140
DB 167237 GCACCTTCCTGCTGCTGTGCGCAGGAATGGAACAGGCTGTGCTGATCAGATCTG 167296

```

```

OY 1141 GCAGCATGGGCGACTGCGTGGATTTCTGCCCAAGACAGCAGAGATGCTGTCTGGCA 1200
DB 167297 GCAGCATGGGCGACTGCGTGGATTTCTGCCCAAGACAGCAGAGATGCTGTCTGGCA 167356
OY 1201 GTGTAAGTCCCGCAGTGTGCTGTGTCCAGAGAGCCACGATGGGAGTCTCTTCTGTGTC 1260
DB 167357 GTGTAAGTCCCGCAGTGTGCTGTGTCCAGAGAGCCACGATGGGAGTCTCTTCTGTGTC 167416
OY 1261 CTCTGCTTCTGTGATCCCTCCACCCCTCTGCTGCTGTGGGCGGAGCCTTTCTCACA 1320
DB 167417 CTCTGCTTCTGTGATCCCTCCACCCCTCTGCTGCTGTGGGCGGAGCCTTTCTCACA 167476
OY 1321 GATCAGTCATTAACCTTAAGACCCCTCATA 1350
DB 167477 GATCAGTCATTAACCTTAAGACCCCTCATA 167506

RESULT 11
BC025174
LOCUS
DEFINITION
MUS musculus, clone MGC:36714 IMAGE:3978387, mRNA, complete cds.
VERSION
BC025174
KEYWORDS
SOURCE
house mouse.
ORGANISM
Mus musculus.
REFERENCE
1 (bases 1 to 1234)
AUTHORS
Strausberg, R.
TITLE
Direct Submission
JOURNAL
Submitted (05-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgaabs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing Center: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amgdbcm.tmc.edu
Gunnarone, P.H., Garcia, A.M., Lu, X., Huylk, S.W., Hale, S.M.,
Yoon, V.S., Kowls, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 61 Row: F Column: 24
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis.
FEATURES
source
Location/Qualifiers
1..1234
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:36714 IMAGE:3978387"
/tissue_type="Mammary tumor. Metallothionien-TGF alpha
model. 10 month old virgin mouse. Taken by biopsy."
/clone_lbp="NCI CGAP_Mam1"
/lab_host="DH10B"
/note="vector: pCMV-Sport6"
112..843
/codon_start=1
/product="Unknown (protein for MGC:36714)"
/protein_id="AAH25174.1"
/db_xref="GI:19263584"

```



```

/product="Similar to DKFZP568B0621 protein"
/protein_id="AAH23068.1"
/db_xref="GI:1860594"
/translation="MRPLLALLLGLVGSPLDNDKTPISLPGQPGILPTPCNNGSG
GLPDRKDRGDRGAPAGAGEGEGEGRPLPEPRGEGPAGCAAGCAACESVPR
RSASAKRSRSEVPAPADTPLPDFRVLINEGNDPTGKTKGCAVYAVANIVAY
RASLQDFLVKNGGSIASFQYEGGMPKASLSCGAMRLKEEDQVAVQGVGDIIGLIT
ASITQDFSLVLSVDFHSSSPVA"
BASE COUNT      258 a      378 g      260 t
ORIGIN          375 c

```

Query Match	50.28;	Score 651.2;	DB 10;	length 1271;
Best Local Similarity	77.9%;	Pred. NO. 1.8e-109;		
Matches 916;	Conservative 0;	Mismatches 238;	Indels 22;	Gaps 6

QY	189	GAGAGGACCCCGGCGTCCGGGGCTCCGGGTCACAGCCCTATGAGGCCACTCTGCTTCG	248
Db	115	GAGAGTTCTCGTAGTCTGAGCTCCGGGGGTACACACCTGAGGGCCACTTCTTGGCCCTTC	174
QY	249	TGCTCTGGGCGCTGCGGGCGGCTGCCCCCAGCTGAGACGACAAAGATCCCCAGCTCT	308
Db	175	TGCTTCTGGGTCTGCTGTGAGCTCTCTCTCTGTGAGGACAAAGATCCCCAGCTCT	234
QY	309	GCCCGGGGACCCCGGCGCTTCCAGGACAGCCCGGGCCACCATGTGACAGCCAGGGCTTGGCG	368
Db	235	GTGCCGGGACGCCCGGCGCTTCCAGGACACGAGTACCATGTGACAGCCAGCCGCTGCTGT	294
QY	369	GCCCGCATGGCCGCGACCGCGCCGAGCCCGGGGCTCCGAGAGAGAAAGGCGAGG	428
Db	295	GCCGTACCGCCGTATGGCTGATGGCCGACGGTGCAACCGGAGCTCCGGGAGAGAAAGGCGAGG	354
QY	429	GCGGAGGCGCGGAGCTGCCGGGACCTCCAGGAGGACCCCGGGCGCGAGAGAGCGGGAC	488
Db	355	GCGGAGAGACGGGACTTACCTGGGGCCAGCTGGGAGACCCCGGGCGCGTGTGAGAGCAGGGC	414
QY	489	CCGCGGGGCCCAACCGGGCTGCGGGGAGTGTCTGCTGCTCCGCGATCCGCTTTCAGCG	548
Db	415	CCATGGGGGCTATCCGGGCTGCGGGGAGTGTCTGTGATCCCCAGATCAGACTTTCAGTG	474
QY	549	CCAAAGCGTCCGAGAGCGGGGTGCTCCGGCTGTGACGCACCCCTTGGCCCTTCAGCGCG	608
Db	475	CCAAAGCATCCGAGAGCGGGGTACTCTCCGCGACGACACCCCTTACCTTTCACACGCTG	534
QY	609	TGCTGTGAACGAGCAGGAGACATTACGAGCGGCTTACCGGCGCAATTACCTGCCAGGTGC	668
Db	535	TGCTGTCAATAGAGCAGGCGCATTTCCGACCCCATCTAGGCAATGTACCTGCCAAGTGC	594
QY	669	CTGGGCTTACTACTTCCCGGCTCCATGCGACACCGCTTACCGGGCCAGCTCGAGTTTGATC	728
Db	595	CTGGGCTTACTACTTCTGCTGCAACGCACTGTCTACCGGGCCAGCTTCAGTTTGATC	654
QY	729	TGCTGAGCAATGGCGAATCATTTGCCCTTCTTCTTCAGTTTTCGCGGGGTGGGCCCAAGC	788
Db	655	TTTGTCAAAMAGGGGACAGTCCATCCCTCTTCTTCCAGTATTTTGGGGGTGGGCCCAAGC	714
QY	789	CAGCTCGCTCTCGGGGGGGCCATGGTGAGGCTGAGGAGCTTGAGAGACAGTGTGGTGC	848
Db	715	CAGCTCGCTCTCAGGGGGGTCCATGGTAAAGGCTAGAACTTGAGGACAGCTGTGGGTGC	774
QY	849	AGTGGGTGGGTGACTACATTGGCACTCTATGCGACATCAAGACAGACAGCACTTCG	908
Db	775	AGTGGGCGGTGGGTGATTACATTGGCACTCTATGCGACATCAAGACAGCAAGTACTTCT	834
QY	909	CCGATTTCTGGGTATCTCCGAGCTGGCGACAGCTCCCGAGCTTCTGTAGTCCCACTGC	968
Db	835	CTGATTTCTGCTCTATTCTGACTGGGACAGCTCCCGAGCTTCTGCTTAAACACAGTGA	894
QY	969	AAAGTACGCTATGCTCTCACTCCTTAGMAAGAGAGGTGTGAGCTGACCAACAGGTATCC	1028
Db	895	ACCGGAGCTGGGCACTTGTCTCTCAGTGGAGGGGTGTGACACTAACCCGGGCGAGCGGATAC	954
QY	1029	AGGAGGCTGGCCCGCTGGAAATTTGTGAATGACTAGGAGGTGGGCTAGACACTCTC	1088

Db	955	CAGGAGGGGCTGGCCCCCTGGGAATATTGTGAATGACTTGGAAAGAGAGGAG-CCACTTCC	1013
QY	1089	CGTCTGCTGCTGGCGAAGAATGGGAACAGTGCTGTCTGCGATCAGCTCGGAGCATG	1148
Db	1014	GGTCCCACTGTGGGCAATGAATGGAGACA--GGCTGTCTGAGATCA---AGACGCTG	1066
QY	1149	GGCGAGTGGCTGAGATTCTGCGCCCAAGACACAGAGAGTGTGCTGTCGGCGCAAGTGTA	1208
Db	1068	GAGCAGTGGCTGGGTCTTCTGCCGAGCATTTA-GAATGCAGTACGCTGCGACGCTGGGT	1128
QY	1209	CCCCAGTGGCTGAGTGCACAGAGCCCAACGCTGGGGTGTCTCTTCTGCTGCTGCTG--	1265
Db	1127	C-----CTGGGCCAGGACTCCAAAGGTGGGATGCTCCATTCCTAGTCCGTGCTCCC	1175
QY	1266	CTTCTCTGAGATCTCCTCCCAACCCCTCTCTGCTCTCTGCGGGGCGGCGCCCTTTTTCACAGATAT	1325
Db	1176	CCCTTAGGTCCCTACATCTCATCTGTGCTGCTCCACAGGCGAGGCGCTTTTTCACAGATCA	1235
QY	1326	CTCATATAACCTAAGAACCTCATTAATAAAAAAAAAAAAAA	1361
Db	1236	CTTAATAAACCTAATAATCTCTCAAAAAAAAAAAAAA	1271

RESULT 13	AP001156	LOCUS	DEFINITION	AP001156	198902 bp	DNA	linear	HTG 30-MAY-2000
			DRAFT SEQUENCE, 42 unordered pieces.					
ACCESSION	AP001156							
VERSION	AP001156.2	GI:8118724						
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT							
SOURCE	Homo sapiens DNA, clone:RP11-657A24.							
ORGANISM	Homo sapiens							

REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

ManuMadaia; Eulheria; Primates; Cacaathini; Homiidae; Homo.
1 (bases 1 to 198902)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 198,902 genomic DNA of 11q23
2 Published Only in Database (2000)
1 (bases 1 to 198902)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (10-FEB-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
Kilaso Univ., 1-15-1 Kilaso, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gs.c.riken.go.jp,
URL:htp://hgp.gs.c.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9824)
On May 31, 2000 this sequence version replaced q1:6997842.

```

----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hnp.gsc.riken.go.jp/
Contact: hatono@gsc.riken.go.jp
----- Project Information
Center project name: HumDraef11
Center clone name: R11-657A24
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 171788 bases at least Q40
Consensus quality: 184443 bases at least Q30
Consensus quality: 191295 bases at least Q20
Insert size: 194802; sum-of-ctrls
Quality coverage: 4.12x in Q20 bases; sum-of-ctrls

```

NOTE: This is a working draft sequence. It currently consists of 42 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps

are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 20627 contig of 20627 bp in length
20728 31817 contig of 11090 bp in length
31918 41131 contig of 9214 bp in length
41232 49745 contig of 8514 bp in length
49846 59457 contig of 9612 bp in length
59558 66804 contig of 7247 bp in length
66905 73970 contig of 7066 bp in length
74071 81007 contig of 6937 bp in length
81108 93250 contig of 6213 bp in length
87421 97965 contig of 4615 bp in length
93351 102941 contig of 5976 bp in length
103042 113391 contig of 4274 bp in length
109118 118783 contig of 5358 bp in length
113492 124241 contig of 5292 bp in length
118884 129766 contig of 5425 bp in length
124342 135473 contig of 5607 bp in length
129867 140585 contig of 5012 bp in length
135574 140686 contig of 3342 bp in length
140686 144027 contig of 3796 bp in length
144128 147923 contig of 4438 bp in length
148024 152461 contig of 2883 bp in length
152562 155444 contig of 3177 bp in length
155545 158721 contig of 3177 bp in length
158722 162803 contig of 4082 bp in length
162904 166121 contig of 3118 bp in length
166222 169851 contig of 3630 bp in length
169952 172048 contig of 2097 bp in length
172149 175332 contig of 3184 bp in length
175433 177914 contig of 2482 bp in length
179796 181604 contig of 1629 bp in length
181705 183951 contig of 2247 bp in length
184052 185927 contig of 1876 bp in length
186028 187870 contig of 1843 bp in length
187971 189591 contig of 1621 bp in length
189692 191169 contig of 1478 bp in length
191270 192527 contig of 1256 bp in length
192628 194288 contig of 1270 bp in length
194369 195658 contig of 100 bp in length
195759 197707 contig of 100 bp in length
197808 198902 contig of 1095 bp in length

Sequence updated (26-May-2000).
NOTE: This is a 'working draft' sequence. It currently
* consists of 42 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 20627 contig of 20627 bp in length
20628 20727: gap of 100 bp
20728 31817: contig of 11090 bp in length
31818 31917: gap of 100 bp
31918 41131: contig of 9214 bp in length
41132 41231: gap of 100 bp
41232 49745: contig of 8514 bp in length
49746 49845: gap of 100 bp
49846 59457: contig of 9612 bp in length
59458 59557: gap of 100 bp
59558 66804: contig of 7247 bp in length
66805 66904: gap of 100 bp
66905 73970: contig of 7066 bp in length
73971 74070: gap of 100 bp
74071 81007: contig of 6937 bp in length
81008 81107: gap of 100 bp
81108 87320: contig of 6213 bp in length
87321 87420: gap of 100 bp
87421 93250: contig of 5830 bp in length
```

FEATURES

source

```
* 93251 93350: gap of 100 bp
* 93351 97965: contig of 4615 bp in length
* 97966 98065: gap of 100 bp
* 98066 102941: contig of 4876 bp in length
* 102942 103041: gap of 100 bp
* 103042 109017: contig of 5976 bp in length
* 109018 109117: gap of 100 bp
* 109118 113391: contig of 4274 bp in length
* 113392 113491: gap of 100 bp
* 113492 118783: contig of 5292 bp in length
* 118784 118883: gap of 100 bp
* 118884 124241: contig of 5358 bp in length
* 124242 124341: gap of 100 bp
* 124342 129766: contig of 5425 bp in length
* 129767 129866: gap of 100 bp
* 129867 135473: contig of 5607 bp in length
* 135474 135573: gap of 100 bp
* 135574 140585: contig of 5012 bp in length
* 140586 140685: gap of 100 bp
* 140686 144027: contig of 3342 bp in length
* 144028 144127: gap of 100 bp
* 144128 147923: contig of 3796 bp in length
* 147924 148023: gap of 100 bp
* 148024 152461: contig of 4438 bp in length
* 152462 152561: gap of 100 bp
* 152562 155444: contig of 2883 bp in length
* 155445 155544: gap of 100 bp
* 155545 158721: contig of 3177 bp in length
* 158722 158821: gap of 100 bp
* 158822 162803: contig of 4082 bp in length
* 162804 163003: gap of 100 bp
* 163004 166121: contig of 3118 bp in length
* 166122 166221: gap of 100 bp
* 166222 169851: contig of 3630 bp in length
* 169852 169951: gap of 100 bp
* 169952 172048: contig of 2097 bp in length
* 172049 172148: gap of 100 bp
* 172149 175332: contig of 3184 bp in length
* 175333 175432: gap of 100 bp
* 175433 177914: contig of 2482 bp in length
* 177915 178014: gap of 100 bp
* 178015 179875: contig of 1861 bp in length
* 179876 179975: gap of 100 bp
* 179976 181604: contig of 1629 bp in length
* 181605 181704: gap of 100 bp
* 181705 183951: contig of 2247 bp in length
* 183952 184051: gap of 100 bp
* 184052 185927: contig of 1876 bp in length
* 185928 186027: gap of 100 bp
* 186028 187870: contig of 1843 bp in length
* 187871 187970: gap of 100 bp
* 187971 189591: contig of 1621 bp in length
* 189592 189691: gap of 100 bp
* 189692 191169: contig of 1478 bp in length
* 191170 191269: gap of 100 bp
* 191270 192527: contig of 1256 bp in length
* 192528 192627: gap of 100 bp
* 192628 194288: contig of 1270 bp in length
* 194289 194388: gap of 100 bp
* 194389 195658: contig of 100 bp in length
* 195659 195758: gap of 100 bp
* 195759 197707: contig of 1949 bp in length
* 197708 197807: gap of 100 bp
* 197808 198902: contig of 1095 bp in length.
```

Location/Qualifiers

1..198902

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q23"

/clone="Rp11-657A24"

misc_feature 1..20627 /note="assembly_fragment"

* arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 1137: contig of 1137 bp in length
 1138 1237: gap of unknown length
 1238 1270: contig of 1483 bp in length
 2721 2820: gap of unknown length
 2821 3984: contig of 1164 bp in length
 3985 4084: gap of unknown length
 4085 5158: contig of 1074 bp in length
 5159 5258: gap of unknown length
 5259 6306: contig of 1048 bp in length
 6307 6407: gap of unknown length
 6407 7524: contig of 1117 bp in length
 7524 7624: gap of unknown length
 7624 8631: contig of 1008 bp in length
 8632 8731: gap of unknown length
 8732 9806: contig of 1075 bp in length
 9807 9906: gap of unknown length
 9907 11155: contig of 1249 bp in length
 11156 11256: gap of unknown length
 11256 12568: contig of 1313 bp in length
 12569 12668: gap of unknown length
 12669 13704: contig of 1036 bp in length
 13705 13804: gap of unknown length
 13805 14931: contig of 1127 bp in length
 14932 15031: gap of unknown length
 15032 16084: contig of 1053 bp in length
 16085 16184: gap of unknown length
 16185 17858: contig of 1674 bp in length
 17859 17958: gap of unknown length
 17959 19035: contig of 1077 bp in length
 19036 19135: gap of unknown length
 19136 20291: contig of 1156 bp in length
 20292 20391: gap of unknown length
 20392 21643: contig of 1252 bp in length
 21644 21743: gap of unknown length
 21744 22846: contig of 1103 bp in length
 22847 22946: gap of unknown length
 22947 24684: contig of 1738 bp in length
 24685 26405: contig of 1621 bp in length
 26406 26505: gap of unknown length
 26506 27716: contig of 1211 bp in length
 27717 27817: gap of unknown length
 27817 28876: contig of 1060 bp in length
 28877 28976: gap of unknown length
 28977 30019: contig of 1043 bp in length
 30020 30120: gap of unknown length
 30121 31550: contig of 1431 bp in length
 31551 31650: gap of unknown length
 31651 33128: contig of 1478 bp in length
 33129 33228: gap of unknown length
 33229 34763: contig of 1535 bp in length
 34764 34863: gap of unknown length
 34864 36404: contig of 1541 bp in length
 36405 36504: gap of unknown length
 36505 37957: contig of 1453 bp in length
 37958 38057: gap of unknown length
 38058 40294: contig of 2236 bp in length
 40295 40393: gap of unknown length
 40394 41698: contig of 1305 bp in length
 41699 41798: gap of unknown length
 41799 43155: contig of 1357 bp in length
 43156 43255: gap of unknown length
 43256 45193: contig of 1938 bp in length
 45194 45293: gap of unknown length
 45294 47317: contig of 1924 bp in length
 47318 49231: contig of 1914 bp in length
 49232 49331: gap of unknown length

49332 51205: contig of 1874 bp in length
 51206 51305: gap of unknown length
 51306 53117: contig of 1812 bp in length
 53118 53217: gap of unknown length
 53218 55343: contig of 2325 bp in length
 55344 55642: gap of unknown length
 55643 57451: contig of 1809 bp in length
 57452 57552: gap of unknown length
 57553 60851: contig of 3300 bp in length
 60852 62718: gap of unknown length
 62719 62819: contig of 1767 bp in length
 62820 64977: gap of unknown length
 64978 65077: contig of 2158 bp in length
 65078 67856: gap of unknown length
 67857 71313: contig of 2779 bp in length
 71314 71413: gap of unknown length
 71414 73659: contig of 3358 bp in length
 73660 73759: gap of unknown length
 73760 75549: contig of 2246 bp in length
 75550 75649: gap of unknown length
 75650 79107: contig of 1790 bp in length
 79108 79206: gap of unknown length
 79207 83551: contig of 3457 bp in length
 83552 83651: gap of unknown length
 83652 86053: contig of 4345 bp in length
 86054 86153: gap of unknown length
 86154 88659: contig of 2402 bp in length
 88660 88759: gap of unknown length
 88760 90078: contig of 2506 bp in length
 90079 90178: gap of unknown length
 90179 92201: contig of 1319 bp in length
 92202 92301: gap of unknown length
 92302 94877: contig of 2023 bp in length
 94878 94977: gap of unknown length
 94978 96946: contig of 2576 bp in length
 96947 97046: gap of unknown length
 97047 99897: contig of 1969 bp in length
 99898 99899: contig of 2851 bp in length

Query Match 37.3% Score 513.4 DB 2: Length 173657;
 Best Local Similarity 78.9% Pred. No. 5.2e-79;
 Matches 717; Conservative 0; Mismatches 171; Indels 21; Gaps 8;

QY 438 CCGGACTGCCGGAGACCTCGAGGAGACCCCGCGGAGAGAGCGGACCGCGGCGC 497
 DB 43765 CAGGACTACCTGAGCGCCAGCTGGGAGAGCGGCGGCGGCGGCGGCGGCGGCGG 43824
 QY 498 CCACCGGGGCGGCGGAGGAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 557
 DB 43825 CTAATCGGCGGCGGCGGAGGAGCTCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 43884
 QY 558 CCGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 617
 DB 43885 CAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 43944
 QY 618 ACGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 677
 DB 43945 ATGACAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 44004
 QY 678 ACTACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 737
 DB 44005 ACTACTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 44064
 QY 738 ATGCGGATCATGCT 797
 DB 44065 ATGCGGATCATGCT 44124
 QY 798 TCTCGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 857
 DB 44125 TCTCGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 44184
 QY 858 TCGGTGCTACATTTGGCATCTATGCCAGCATCAAGACAGACACCTTCTCCGATTTTC 917

```

Db      44185 TGGGTATTACATTTGGCATCTATGCCAGATCAAAACAGACATACCTTCTGTGATTTC 44244
QY      918 TGGTACTACCGACAGGACAGCTCCCAAGCTTTGCTTAAGTCCCACTGCAAGTGAGC 977
Db      44245 TGGTCTATTCTACTGCGACAGCTCCCAAGCTTCTGCTTAATAATACAGTGAACCCGAGC 44304
QY      978 TATGCTCTACTCTAGAGAGAGGTGTGAGCTGACCAACAGGTATCATCCAGAGAGGCT 1037
Db      44305 T--GGCACTTGTCTCTAG--TGAAGGTGTGTACACTGACAGCGGCAATACAGAGGCT 44361
QY      1038 GGGCCCGCTGGAATATTTGTAATCACTAGAGAGGTGGGTAGAGACTCTCCCTGCT 1097
Db      44362 GG--CCCCCTGACATATGTGATGATAGAAAGAGAGG--ACCCACCTCTGTCTCCCT 44418
QY      1098 GCTGGACGAAGATGGAAGAGAGGTGTGCTGCAATGCTGCGAGCATATGGGAGAGTG 1157
Db      44419 CCTGGCAATGMAATAG--TCAAGAGCTGTCTGAGAGTCAAGACAGCGGTGAGAGAGTTG 44476
QY      1158 CTGATTTTGGCCCAACAGAGAGAGTGTCTGTCTGCAAGTGTAAAGTCCCAAGTT 1217
Db      44477 CTGGGTTTCTGCGACAGACTTGTAAATGACAGAG--GCTGGCAGCTGT-----GGA 44525
QY      1218 GCTGTGTGACAGAGCCAGCGAGGTGGGTGCTCTCTGCTGCTGCTGCTGCTGCTGATC 1277
Db      44526 ATCTGTGCTGACAGCTCAAGATGGATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 44585
QY      1278 CTGCCCCAGC--CGCTGCTGCTCTCTGAGGCGGCGGCTTTTCTAGAGATCAATCAATAC 1335
Db      44586 TTGTGACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 44645
QY      1336 CTAGAGACC 1344
Db      44646 CTAGCAATCC 44654

RESULT 15
ARI38195
LOCUS      ARI38195
DEFINITION Sequence 10 from patent US 6197930.
ACCESSION  ARI38195
VERSION    ARI38195.1 GI:14479704
KEYWORDS
SOURCE
ORGANISM   Unknown.
REFERENCE  1 (bases 1 to 729)
AUTHORS   Sheppard, P.O. and Humes, J.M.
TITLE     Adipocyte-specific protein homologs
JOURNAL   Patent: US 6197930-A 10 06-MAR-2001;
FEATURES
source     1..729
            /organism="unknown"
BASE COUNT 77 a 101 c 160 g 81 t 310 others
ORIGIN

```

Query Match 36.6%; Score 504; DB 6; Length 729;
 Best Local Similarity 57.6%; Pred. No. 3.4e-77;
 Matches 419; Conservative 145; Mismatches 164; Indels 0; Gaps 0;

Search completed: June 22, 2003, 04:43:30
 Job time : 3606 secs

```

Db      181 GCNCCNGNGARNAARGGAGGAGNGGNGMNCNCGNGVNTNCCNCGNCGNNGNAGYCCN 240
QY      467 GGGCCGCGAGAGAGGCGGAGACCGCGGGGCGCCACCGGGGCTCCCGGGAGAGTCTGGTG 526
Db      241 GGNCCNMGNGNGARNGNGNCCNCGNCGNCAACNCGNCGNCGNCGNCGNCGNCGNCGNCGN 300
QY      527 CCTCCGGGATCCGCTTACACCGCCCAAGCCCTCCGAGAGCGGGGTGCTCCCGCTGTGAC 586
Db      301 CCNCCNMGNMSNCCNTTYSNCGNMAARMGNWSNGARMSNMGNTNCCNCCNCGNAGY 360
QY      587 GCACCCCTTGCCTTGCACCGCGGTGTGTGAACGAGAGGAGGACATTTAGAGCCGCTCAC 646
Db      361 GCNCCNNTNCCNTTGTAGYMGNGNTNNGTNAAYGANCARGNCATTAAGAYGACNCTNACN 420
QY      647 GGCAGTTACCTGCCAGAGTGCCTGGGGTCTACTACTTGGCGTCCATGACCAAGCTGTAC 706
Db      421 GGNAAATTACNTGYCARGTNCNCGNNGTNTAYTAYTGYCCNGTNCAGCAACNCTNTAY 480
QY      707 CGGCGCAGCCTGACGTTGATCTGTGTGAAGATGGCGAATCCATTGCTCTTTCTTCAG 766
Db      481 MGNCGMWSNVTNCAATTTAYATNTNAAARAYGNGARMSNATHGCMWSNTTYYTCAR 540
QY      767 TTTTGGGGGGTGGCCCAAGCCAGCTGCTCTGGGGGGGCGCATGTGAGGCTGGAG 826
Db      541 TTTTGGGGGGTGGCCCAARCCNCCNMSNVTNMSNCGNCGNCGNCGNCGNCGNCGN 600
QY      827 CCTGAGGACCAAGTGTGGTGCAGGTGGGTGGGTGATGATGATGATGATGATGATGATGAT 886
Db      601 CCGARGATCARGTGTGTCNCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 660
QY      887 ATCAAGACAGACAGACTTCTCCGATTTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 946
Db      661 ATHAARACGATWSNACNTTYSNCGNNTTYYTNTNAYNSNAGYTGCGAYMSNCCN 720
QY      947 GTCTTTCG 954
Db      721 GTNTTTCG 728

```

PR	29-APR-1998;	98US-0083500
PR	22-MAY-1998;	98US-0086414

PR 10-JUN-1998: 98US-0086742.
 PR 10-NOV-1998: 98US-0107783.
 XX (GETH) GENENTECH INC.
 XX Botstein D, Goddard A, Gurney AL, Hillan KJ, Lawrence DA;
 PI Roy MA, Wood WI;
 XX MPI: 1999-430385/36.
 DR P-PSDB; AAY06481.
 XX
 PT Antibody against proteins expressed in neoplastic cells, useful for
 PT tumor diagnosis and treatment
 XX
 PS Example 1; Fig 9; 162pp; English.
 XX
 CC This is the nucleotide sequence of cDNA clone DNA40592 (ATCC 209492)
 CC coding for human PRO344 (UNQ303) (see AAY06481). The clone was
 CC isolated from a foetal kidney library. Amplification of DNA40592
 CC was observed in primary lung tumours and in primary colon tumours,
 CC suggesting an association with tumour formation or growth.
 CC Antagonists (e.g. antibodies) directed against PRO344 may have
 CC utility in cancer therapy. The invention identifies 14 genes (see
 CC AAY07254-67) that are amplified in the genome of tumour cells. Such
 CC amplification is expected to be associated with overexpression of
 CC the gene product and to contribute to tumorigenesis. The encoded
 CC proteins (see AAY06477-90) may be useful targets for the diagnosis
 CC and/or treatment (including prevention) of certain cancers, and may
 CC act as predictors of the prognosis of tumour treatment.
 XX
 SO Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other:
 Query Match 100.0%; Score 1377; DB 20; Length 1377;
 Best Local Similarity 100.0%; Pred. No. 1.7e-244;
 Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACTGTTCTCTTGGAGTCTGGAAGAGAAAGCGAGCGGAGGAGGAAACCGAGC 60
 DB 1 GACTGTTCTCTTGGAGTCTGGAAGAGAAAGCGAGCGGAGGAGGAAACCGAGC 60
 QY 61 TGGGGTACGCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 DB 61 TGGGGTACGCGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 QY 121 CCAACTGAGAGGCTCGAGTACGAGAGGCGCCGCAAGAGAGGAGGAGGAGGAGGAGG 180
 DB 121 CCAACTGAGAGGCTCGAGTACGAGAGGCGCCGCAAGAGAGGAGGAGGAGGAGGAGG 180
 QY 181 GGGACTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 DB 181 GGGACTCGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
 QY 241 GGTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 DB 241 GGTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 QY 301 CAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 DB 301 CAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 QY 361 CTTGCGCGGCGCGAGTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGG 420
 DB 361 CTTGCGCGGCGCGAGTGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGG 420
 QY 421 AGCGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGG 480
 DB 421 AGCGGAGGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGG 480
 QY 481 GGGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGG 540
 DB 481 GGGGAGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGG 540
 QY 541 CTTGAGGCGCAAGCGCTCGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGG 600

DB 541 CTTGAGGCGCAAGCGCTCGAGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGG 600
 QY 601 CGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 DB 601 CGACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 660
 QY 661 CCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 DB 661 CCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
 QY 721 GTTGTATCTGTGTAAGATGCGAATGCGAATGCGAATGCGAATGCGAATGCGAATG 780
 DB 721 GTTGTATCTGTGTAAGATGCGAATGCGAATGCGAATGCGAATGCGAATGCGAATG 780
 QY 781 GCCCAAGCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 DB 781 GCCCAAGCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 QY 841 GTGGGTGCAAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 900
 DB 841 GTGGGTGCAAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 900
 QY 901 CACCTTCTCGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 DB 901 CACCTTCTCGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 QY 961 CCCACTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
 DB 961 CCCACTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1020
 QY 1021 GGTGATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
 DB 1021 GGTGATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
 QY 1081 GCACCTCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 DB 1081 GCACCTCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
 QY 1141 GCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
 DB 1141 GCAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1200
 QY 1201 GTGTAAGTCCCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1260
 DB 1201 GTGTAAGTCCCGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1260
 QY 1261 CTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
 DB 1261 CTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
 QY 1321 GATCAGTCAATTAACCTAAGAAACCTCAATTAAGAAACCTCAATTAAGAAACCTCA 1377
 DB 1321 GATCAGTCAATTAACCTAAGAAACCTCAATTAAGAAACCTCAATTAAGAAACCTCA 1377
 RESULT 2
 ID AAX80052 standard: cDNA; 1377 BP.
 XX AAX80052:
 AC AAX80052:
 XX
 DT 12-AUG-1999 (first entry)
 XX
 DE Human PRO344 nucleotide sequence.
 XX
 KW Human: PRO protein; tumour necrosis factor family; TNF; cytokine;
 KW secreted protein; transmembrane protein; inflammation disorder; s9.
 OS Homo sapiens.
 XX
 PN WO9928462-A2.
 XX

```

PD 10-JUN-1999.
XX
XX
PF 01-DEC-1998; 98WO-US25108.
XX
XX
PR 25-FEB-1998; 98US-0075945.
PR 03-DEC-1997; 97US-0067411.
PR 11-DEC-1997; 97US-0069278.
PR 11-DEC-1997; 97US-0069334.
PR 11-DEC-1997; 97US-0069335.
PR 12-DEC-1997; 97US-0069425.
PR 16-DEC-1997; 97US-0069696.
PR 16-DEC-1997; 97US-0069702.
PR 17-DEC-1997; 97US-0069870.
PR 17-DEC-1997; 97US-0069873.
PR 18-DEC-1997; 97US-0068017.
PR 05-JAN-1998; 98US-0070440.
PR 09-FEB-1998; 98US-0074086.
PR 09-FEB-1998; 98US-0074092.
XX
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Chen J, Goddard A, Gurney AL, Wood WI;
PI Yuan J;
XX
XX WPI: 1999-371118/31.
DR P-PSDB; AAY17827.
XX
XX
XX Nucleic acids encoding PRO secreted and transmembrane proteins
XX
XX Claim 2; Fig 20; 123pp; English.
XX
XX
XX The present invention describes nucleic acids encoding PRO secreted and
XX transmembrane proteins used therapeutically. The PRO proteins have
XX cytosolic, anti-inflammatory, anti-proliferative and immunosuppressive
XX activity. The proteins and polynucleotides can be used in therapy,
XX identification of homologues, raising antibodies and design of probes
XX and primers. They can be used in a range of diseases related to proteins
XX that they have homology with, e.g. a PRO protein having homology to
XX complement proteins may be used in inflammatory responses.
XX
XX
XX Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other;
XX
XX
XX Query Match 100.0%; Score 1377; DB 20; Length 1377;
XX Best Local Similarity 100.0%; Pred No.1.7e-244;
XX Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX
XX 1 GACTAGTCTCTTGGAGTCTGGAGAGCAAGCCGAGCCGCGAGAGCAACCGAGAC 60
XX |||||
XX 1 GACTAGTCTCTTGGAGTCTGGAGAGCAAGCAAGCCGAGCCGCGAGAGCAACCGAGAC 60
XX
XX 61 TGGGCTGACGCGAGGCGAGGCGGCGCCCTGCGCGGCGAGCAAGCGCGGCGCTTGAGACCA 120
XX |||||
XX 61 TGGGCTGACGCGAGGCGAGGCGGCGCCCTGCGCGGCGAGCAAGCGCGGCGCTTGAGACCA 120
XX
XX 121 CCAACTGAGAGGCTCCGAGTAGCAGACGCCGCCCAAGAGAGCCATCGGGAGCGGAGAGG 180
XX |||||
XX 121 CCAACTGAGAGGCTCCGAGTAGCAGAGCGCCGCCCAAGAGAGCCATCGGGAGCGGAGAGG 180
XX
XX 181 GGGACTGGGAAGAGACCCCGGCGCTCCGGGCTCCCGGTGCCAGGCGCTATGAGGCCACTCT 240
XX |||||
XX 181 GGGACTGGGAAGAGACCCCGGCGCTCCGGGCTCCCGGTGCCAGGCGCTATGAGGCCACTCT 240
XX
XX 241 CGTCTGCTGCTCTTGGGCTGCGGCGCGGCTGCCCGCACTGGAGAGCAACAAGATCTCC 300
XX |||||
XX 241 CGTCTGCTGCTCTTGGGCTGCGGCGCGGCTGCCCGCACTGGAGAGCAACAAGATCTCC 300
XX
XX 301 CAGGCTTGCCCGGGGCGACCCCGGCGCTTCCAGGACGCGCGGCGCACCATGAGCGACGAG 360
XX |||||
XX 301 CAGGCTTGCCCGGGGCGACCCCGGCGCTTCCAGGACGCGCGGCGCACCATGAGCGACGAG 360
XX
XX 361 CTTCGCGGCGCGGATGCGCGACGCGCGGAGCGGCGCGGCGCTCCGGGAGAGAA 420
XX |||||

```

Db	361	CTTGCCGGGCGCGCATGGCCCGGCACAGGCCGCGCCCGGGGCTCCGGAGAGAA	420
OY	421	AGGCGAGGGCGGGAGAGCCGGGACTCCGGGACCTCGAGGGGACCCCGGCGGAGAGA	480
Db	421	AGGCGAGGGCGGGAGAGCCGGGACTCCGGGACCTCGAGGGGACCCCGGCGGAGAGA	480
OY	481	GGCGGAGCCCGGGGGCCACACCGGGGCTGCGCGGGGAGTGGTCGGTGCCTCCGAGTCCG	540
Db	481	GGCGGAGCCCGGGGGCCACACCGGGGCTGCGCGGGGAGTGGTCGGTGCCTCCGAGTCCG	540
OY	541	CTTGAGGGCCCAAGCGGCTCCGAGAGCCGGGGTCCGCGCGCTGAGCAGACCTTGGCCCTT	600
Db	541	CTTGAGGGCCCAAGCGGCTCCGAGAGCCGGGGTCCGCGCGCTGAGCAGACCTTGGCCCTT	600
OY	601	CGACCGCGCTGCTGGTGAACGAGCAGGAGACATTTAGCAGCGCGCTACCGGCAATGTACCTG	660
Db	601	CGACCGCGCTGCTGGTGAACGAGCAGGAGACATTTAGCAGCGCGCTACCGGCAATGTACCTG	660
OY	661	CCAGGTCCCGGGGGTCTACTACTTCCGCGCTGCATGCGCCACCGCTACCGGGCCAGCGCTGA	720
Db	661	CCAGGTCCCGGGGGTCTACTACTTCCGCGCTGCATGCGCCACCGCTACCGGGCCAGCGCTGA	720
OY	721	GTTTGAATCTGGTGAAGAAATGGCCGAATCCATTTGCGCTTTTCTTCCAGTTTTCGGGGGGT	780
Db	721	GTTTGAATCTGGTGAAGAAATGGCCGAATCCATTTGCGCTTTTCTTCCAGTTTTCGGGGGGT	780
OY	781	GGCCAAAGCCAGCCCTCCCTCTCGGGGGGGGCCATGCTGAGCGGTGAGACCTGAGACCAAGT	840
Db	781	GGCCAAAGCCAGCCCTCCCTCTCGGGGGGGGCCATGCTGAGCGGTGAGACCTGAGACCAAGT	840
OY	841	GTGGGTCGAGGTGGGGTGGTGACATCAATTTGGCATTTTGGCAGACATCAAGACAGACAG	900
Db	841	GTGGGTCGAGGTGGGGTGGTGACATCAATTTGGCATTTTGGCAGACATCAAGACAGACAG	900
OY	901	CACCTTCCTCCGGAATTTCTGGTGTACTCCGACTGGCAGAGCTCCCGAGTCTTGTCTTAAGT	960
Db	901	CACCTTCCTCCGGAATTTCTGGTGTACTCCGACTGGCAGAGCTCCCGAGTCTTGTCTTAAGT	960
OY	961	CCCACTGCAAACTGAGCTCATGCTCTCACTCCTTAGAAGAGGGGTGAGGGTGACAAACA	1020
Db	961	CCCACTGCAAACTGAGCTCATGCTCTCACTCCTTAGAAGAGGGGTGAGGGTGACAAACA	1020
OY	1021	GGTCAATCCAGAGGGGCTGGCCCCCTGGCAATATTTGTGAATACTAGAGGAGTGGGTAGA	1080
Db	1021	GGTCAATCCAGAGGGGCTGGCCCCCTGGCAATATTTGTGAATACTAGAGGAGTGGGTAGA	1080
OY	1081	GCACCTCTCCGTCCTGCTGCTGGCAAGAAATGGGAAACATGGCTTCTGCGCATCAGGCTG	1140
Db	1081	GCACCTCTCCGTCCTGCTGCTGGCAAGAAATGGGAAACATGGCTTCTGCGCATCAGGCTG	1140
OY	1141	GCAGATGGGGCAATGGTGTGATTTCTGCGCCAGACAGAGAGAGTGTGCTGTGCAGAA	1200
Db	1141	GCAGATGGGGCAATGGTGTGATTTCTGCGCCAGACAGAGAGAGTGTGCTGTGCAGAA	1200
OY	1201	GTGTAAGTCCCGCAGTTGCTGTGGTCCAGAGACCCAGCGAGGGGTGCTCTTCCTCGTGC	1260
Db	1201	GTGTAAGTCCCGCAGTTGCTGTGGTCCAGAGACCCAGCGAGGGGTGCTCTTCCTCGTGC	1260
OY	1261	CTTGCTTCTCTGGATCTCCCAACCCCTCTGCTGCTGGGGCGGCGCCCTTTTCTCAGA	1320
Db	1261	CTTGCTTCTCTGGATCTCCCAACCCCTCTGCTGCTGGGGCGGCGCCCTTTTCTCAGA	1320
OY	1321	GATCACTCAATTAACCTTAAGAACCCCTCATTAATAAAAAAAAAAAAAAAAAAAAAA	1377
Db	1321	GATCACTCAATTAACCTTAAGAACCCCTCATTAATAAAAAAAAAAAAAAAAAAAAAA	1377
RESULT 3			
AD01241			
ID	AD01241	standard; cDNA; 1377 BP.	
XX			
AC	AD01241;		

sequence 13// BP; 251 A; 422 C; 472 G; 232 T; 0 other,

201 CCCAC GCAAGAG GAGCTCAT GCTCTT CACTCCTAG AAGGAGG GTGTGAG GCTGACA ACCA 1020

Db 841 GTGGGTGAGGTGGTGTGGTGTGACATGATGATCCAGCATCAAGACAGAG 900
 QY 901 CACCTTCCGGATTTCTGTGTACTCCAGATGGACAGATCCCGAGTCTTGTAGTG 960
 Db 901 CACCTTCCGGATTTCTGTGTACTCCAGATGGACAGATCCCGAGTCTTGTAGTG 960
 QY 961 CCCAGTCAAAAGTAGTCTACTCTCTACTCTCTAGAGAGAGGTGTAGAGTCAACCA 1020
 Db 961 CCCAGTCAAAAGTAGTCTACTCTCTACTCTCTAGAGAGAGGTGTAGAGTCAACCA 1020
 QY 1021 GGTATCAGAGAGAGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1080
 Db 1021 GGTATCAGAGAGAGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 1080
 QY 1081 GCACTCTCCGCT 1140
 Db 1081 GCACTCTCCGCT 1140
 QY 1141 GCAGCATGGGAGAGT 1200
 Db 1141 GCAGCATGGGAGAGT 1200
 QY 1201 GTGTAGTCCCGCATGCTGT 1260
 Db 1201 GTGTAGTCCCGCATGCTGT 1260
 QY 1261 CTCTGCTCTCTGT 1320
 Db 1261 CTCTGCTCTCTGT 1320
 QY 1321 GATCCTCAATTAACCTTAAGAACCTCTATMAAAAAAAAAAAAAAAAAAAAAA 1377
 Db 1321 GATCCTCAATTAACCTTAAGAACCTCTATMAAAAAAAAAAAAAAAAAAAAAA 1377
 RESULT 6
 AAS21424
 ID AAS21424 standard; cDNA: 1377 BP.
 AC AAS21424:
 XX 24-OCT-2001 (first entry)
 XX Human cDNA sequence encoding for PRO344 polypeptide.
 XX Human secretory and transmembrane; PRO; mammalian; cancer; lung;
 KW breast; prostate; cervical; tumour necrosis factor-alpha; TNF-alpha;
 KW catilage; ear; proliferation; glucose; free fatty acid; skeletal muscle;
 KW adipocyte; A-peptide; factor VIIA; gene therapy; ss.
 OS Homo sapiens.
 XX
 XX MO200140466-A2.
 XX
 PD 07-JUN-2001.
 PF 01-DEC-2000; 2000MO-US32678.
 XX
 PR 01-DEC-1999; 99MO-US28301.
 PR 02-DEC-1999; 99MO-US28634.
 PR 02-DEC-1999; 99MO-US28531.
 PR 02-DEC-1999; 99MO-US28564.
 PR 09-DEC-1999; 99MO-US28565.
 PR 16-DEC-1999; 99MO-US30095.
 PR 20-DEC-1999; 99MO-US30911.
 PR 20-DEC-1999; 99MO-US30999.
 PR 30-DEC-1999; 99MO-US31243.
 PR 06-JAN-2000; 2000MO-US00277.
 PR 11-FEB-2000; 2000MO-US00376.
 PR 18-FEB-2000; 2000MO-US03565.
 PR 18-FEB-2000; 2000MO-US04341.
 PR 18-FEB-2000; 2000MO-US04342.

PR 22-FEB-2000; 2000MO-US04414.
 PR 24-FEB-2000; 2000MO-US04914.
 PR 24-FEB-2000; 2000MO-US05004.
 PR 01-MAR-2000; 2000MO-US05601.
 PR 20-MAR-2000; 2000MO-US07377.
 PR 21-MAR-2000; 2000MO-US07532.
 PR 30-MAR-2000; 2000MO-US08439.
 PR 17-MAY-2000; 2000MO-US13705.
 PR 22-MAY-2000; 2000MO-US14042.
 PR 30-MAY-2000; 2000MO-US14941.
 PR 02-JUN-2000; 2000MO-US15264.
 PR 10-NOV-2000; 2000MO-US30873.
 XX
 PA (GERTH) GENENTECH INC.
 XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W,
 PI Gerlitsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 DR WPI: 2001-408281/43.
 DR P-PSDB; AA012352.
 XX
 PT Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical
 XX
 PS Claim 3; Fig 361; 813pp; English.
 XX
 CC AAS21244-AAS21518 encode for novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PMCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 CC
 XX Sequence 1377 BP; 251 A; 422 C; 472 G; 232 T; 0 other:
 SO
 Query Match 100.0%; Score 1377; DB 22; Length 1377;
 Best Local Similarity 100.0%; Pred. No. 1,7e-244;
 Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GACTACTTCTCTTGAAGTCTGGAGAGAGAAAGCGAGCCCGAGAGGAGCGAAGACAGAG 60
 Db 1 GACTACTTCTCTTGAAGTCTGGAGAGAGAAAGCGAGCCCGAGAGGAGCGAAGACAGAG 60
 QY 61 TGGGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 Db 61 TGGGGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
 QY 121 CCAACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 Db 121 CCAACTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
 QY 181 GGGAGTGGAG 240
 Db 181 GGGAGTGGAG 240
 QY 241 CGTCTGCT 300

protein zs1939 (AA097984). Claimed polynucleotides comprise nucleotides 243-962, 252-962, 285-482, 285-485, 285-488, 285-491, 285-491, 491-926 of this isolated polynucleotide. Novel zs1939 polypeptides were initially identified by querying the EST database for secretory signal sequences characterised by an upstream methionine start site, a hydrophobic region of approximately 13 amino acids and a cleavage site. A single EST sequence was discovered, and the novel polypeptide encoded by this sequence was cloned. A full-length clone was obtained from a lung tissue library. The zs1939 gene maps to the 11q23.3 region. Expression vectors, cultured cells and a method of producing zs1939 polypeptide are claimed, as well as probes and primers (useful in diagnostic applications), and a method for modulating free fatty acid metabolism by administering a zs1939 polypeptide. The zs1939 polypeptide may also be used in organ preservation, for cytopreservation, for surgical pretreatment to prevent injury due to ischaemia and/or inflammation, and as an antimicrobial agent, promoting lysis or phagocytosis of infectious agents.

Sequence 1347 BP; 235 A; 421 C; 465 G; 226 T; 0 other;

[illegible]

QY 1300 GGGCGCGCCCTTTCTGACAGATCAGTCAATAAACCTAGAACCTCTATAAAAAA 1359
DB 1271 GGGCGCGCCCTTTCTGACAGATCAGTCAATAAACCTAGAACCTCTATAAAAAA 1330
QY 1360 AAAAAA 1366
DB 1331 AAAAAA 1337

RESULT 9
AAFA4970
ID AAFA4970 standard; cDNA: 1338 BP.
AAFA4970;
AC AAFA4970;
XX
XX
XX 28-MAR-2001 (first entry)
DE Human TANGO 253 coding sequence SEQ ID NO: 1.
XX
XX
XX Human; mouse; secreted protein: TANGO253; TANGO 257; TANGO 281.
KM INTERCEPT 258; coronary disorder; olfactory disorder;
KM neurological disorder; pulmonary disorder; immunological disorder;
KW developmental disorder; kidney disorder; ss.
XX Homo sapiens.
XX
XX MO200078808-A1.
XX
XX 28-DEC-2000.
XX
XX 19-JUN-2000; 2000MO-US16883.
XX
XX 18-JUN-1999; 99US-0336536.
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Leiby KR, McKay C, Bosone S;
XX WPI; 2001-050109/06.
XX
XX
XX New nucleic acids for treating diseases and disorders, e.g.
PT atherosclerosis, infection, autoimmune diseases, obesity, ear
PT disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT sclerosis and asthma.
XX
XX
XX Claim 1; Page 210; 332pp; English.
XX
XX The present invention provides the protein and coding sequences of the
CC human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC coronary, pulmonary, olfactory, immunological, neurological,
CC developmental and kidney disorders.
XX
XX Sequence 1338 BP; 229 A; 423 C; 460 G; 226 T; 0 other;

Query Match 94.2%; Score 1296.6; DB 22; Length 1338;
Best Local Similarity 99.6%; Pred. No. 1e-229;
Matches 1310; Conservative 0; Mismatches 4; Indels 1; Gaps 1;

QY 54 CCAGACTGGGTGACGCGCAGGAGGAGGCGGCTGCGCGGAGAGCGGGGCTGG 113
DB 15 CCGGACTGGGTGACGCGCAGGAGGAGGCGGCTGCGCGGAGAGCGGGGCTGG 74
QY 114 AGACACCACTGGAGGCTCCGAGTAGAGGAGCGCCCGCAAGAGGCGCATCGGGAGCC 173
DB 75 AGACACCACTGGAGGCTCCGAGTAGAGGAGCGCCCGCAAGAGGCGCATCGGGAGCC 134
QY 174 GGGAGGGGGGCTCGAGAGAGACCCCGGCTCCGGGCTCCGGTGCAGAGGCTATAGAGC 233
DB 135 GGGAGGGGGGCTCGAGAGAGACCCCGGCTCCGGGCTCCGGTGCAGAGGCTATAGAGC 194
QY 234 CACTCTCTGCTCTCTCTCTCTGCGGCTGCGGCGGCTGCGCCCACTGAGAGACACA 293

DB 195 CACTCTCTGCTCTCTCTCTCTGCGGCTGCGGCGGCTGCGCCCACTGAGAGACACA 254
QY 294 AGATCCCGAGGCTCTCCCGGGGAGACCCCGGCTTTCAGAGCAGCGGGGAGCAGATGCA 353
DB 255 AGATCCCGAGGCTCTCCCGGGGAGACCCCGGCTTTCAGAGCAGCGGGGAGCAGATGCA 314
QY 354 GCCAGGGCTTGGCCGGGCGCGATGGCGCGGAGCGCGGAGCGGCGGGGCTCCGG 413
DB 315 GCCAGGGCTTGGCCGGGCGCGATGGCGCGGAGCGCGGAGCGGCGGGGCTCCGG 374
QY 414 GAGAGAAAGGAGGCGGGGAGAGCGCGGAGCTGCGGGAGCTGAGAGGAGCCCGGCGCG 473
DB 375 GAGAGAAAGGAGGCGGGGAGAGG -CGGGAGCTGCGGGAGCTGAGAGGAGCCCGGCGCG 433
QY 474 GAGAGAGGCGGGGAGCGCGGCGGAGCGCGGAGCTGCGGGAGCTGAGAGGAGCCCGGCGCG 533
DB 434 GAGAGAGGCGGGGAGCGCGGCGGAGCGCGGAGCTGCGGGAGCTGAGAGGAGCCCGGCGCG 493
QY 534 GATCGGCTTTCAGCGCGGAGAGCGCTCCGAGAGCGCGGCTCCGCTGAGAGCAGCCT 593
DB 494 GATCGGCTTTCAGCGCGGAGAGCGCTCCGAGAGCGCGGCTCCGCTGAGAGCAGCCT 553
QY 594 TGCCCTTTCAGCGCGGCTTTCAGAGAGGAGGAGCTTTCAGAGCGCGGCTTTCAGAGCAGT 653
DB 554 TGCCCTTTCAGCGCGGCTTTCAGAGAGGAGGAGCTTTCAGAGCGCGGCTTTCAGAGCAGT 613
QY 654 TGCCCTTTCAGCGCGGCTTTCAGAGAGGAGGAGCTTTCAGAGCGCGGCTTTCAGAGCAGT 713
DB 614 TGCCCTTTCAGCGCGGCTTTCAGAGAGGAGGAGCTTTCAGAGCGCGGCTTTCAGAGCAGT 673
QY 714 GCTGCGAGTTTGTCTGTGTAAGATGGGAGATTCATTCGCTCTTTCAGAGTTTTCG 773
DB 674 GCTGCGAGTTTGTCTGTGTAAGATGGGAGATTCATTCGCTCTTTCAGAGTTTTCG 733
QY 774 GGGGAGGCGGAGAGCGAGCGCTCTCGGGGGGAGAGGAGGAGGAGGAGGAGGAGG 833
DB 734 GGGGAGGCGGAGAGCGAGCGCTCTCGGGGGGAGAGGAGGAGGAGGAGGAGGAGG 793
QY 834 ACCAAGTGTGGTGCAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG 893
DB 794 ACCAAGTGTGGTGCAGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGGAGTGGG 853
QY 894 CAGACAGACCTTCTCCGATTTCTGTGCTAGCTGCGAGCTGCGAGCTCCGAGCTTTCG 953
DB 854 CAGACAGACCTTCTCCGATTTCTGTGCTAGCTGCGAGCTGCGAGCTCCGAGCTTTCG 913
QY 954 CTTAGTGGCCAGTGGCAAGTGGATCTCTACCTCTAGAGAGGAGTGGAGGCTG 1013
DB 914 CTTAGTGGCCAGTGGCAAGTGGATCTCTACCTCTAGAGAGGAGTGGAGGCTG 973
QY 1014 ACAACAGGTCATCCAGAGAGGCGTGGCCCGCTGGAATATTGTGAATAGTAAAGAGGAGT 1073
DB 974 ACAACAGGTCATCCAGAGAGGCGTGGCCCGCTGGAATATTGTGAATAGTAAAGAGGAGT 1033
QY 1074 GGGTGGAGACTCTCCGCTCTCTCTCTCTGCGCAAGAGTGGAGTGGAGTGGAGTGG 1133
DB 1034 GGGTGGAGACTCTCCGCTCTCTCTCTCTGCGCAAGAGTGGAGTGGAGTGGAGTGG 1093
QY 1134 AGCTTGGCAGAGTGGGAGTGGGCTGATTTTCGCGCCAGAGAGGAGTGGGCTG 1193
DB 1094 AGCTTGGCAGAGTGGGAGTGGGCTGATTTTCGCGCCAGAGAGGAGTGGGCTG 1153
QY 1194 CTGGCAATGTAAATCCCGCAGTTGCTGTGTCAGAGAGCCAGAGTGGGCTGCTCTT 1253
DB 1154 CTGGCAATGTAAATCCCGCAGTTGCTGTGTCAGAGAGCCAGAGTGGGCTGCTCTT 1213
QY 1254 CTTGCT 1313
DB 1214 CTTGCT 1273
QY 1314 TCTGAGAGATCT 1368

AC	AFA4999;
XX	
DT	28-MAR-2001 (first entry)
XX	
DE	
XX	Human secreted protein related coding sequence SFG ID NO: 105.
KW	Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;
KM	INTERCEPT 256; coronary disorder; olfactory disorder;
KW	neurological disorder; pulmonary disorder; immunological disorder;
KX	developmental disorder; kidney disorder; ss.
OS	Homo sapiens.
XX	
PN	WO200078808-A1.
XX	
PD	28-DEC-2000.
XX	
PF	19-JUN-2000; 2000MO-US16083.
XX	
PR	18-JUN-1999; 99US-0336536.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
PI	Lelby KR, McKay C, Bossone S;
XX	WPI: 2001-050109/06.
DR	New nucleic acids for treating diseases and disorders, e.g.
XX	atherosclerosis, infection, autoimmune diseases, obesity, ear
PT	disorders, brain disorders, tumors, diabetes, arthritis, multiple
PT	sclerosis and asthma -
XX	
PS	Disclosure: Page 272; 332pp; English.
CC	The present invention provides the protein and coding sequences of the
CC	human and murine secreted or transmembrane proteins TANGO 253, TANGO 257,
CC	TANGO 281 and INTERCEPT 258. These are useful in the treatment of
CC	coronary, pulmonary, olfactory, immunological, neurological,
CC	developmental and kidney disorders.
SQ	Sequence 1338 BP: 229 A; 422 C; 460 G; 227 T; 0 other;
Query Match	94.0%; Score 1295; DB 22; Length 1338;
Best Local Similarity	99.5%; Pred. No. 2e-229;
Matches 1309; Conservative	0; Mismatches 5; Indels 1; Gaps 1.
OY	54 CCAGGACTGGGTGCATGCCGACGGCGACGGGCGCTTGCCGGGAGAAGCGCGGGGCTTGG 113
DG	15 CGGGACTGGGTGCATGCCGACGGCGACGGGCGCTTGCCGGGAGAAGCGCGGGGCTTGG 74
OY	114 AGCACCACCACTGTGAGGAGTCGCGAGTAGCGAGCCCCCGAAGAGGCCATTGCGGAGCC 173
DG	75 ACACACCACTGTGAGGAGTCGCGAGTAGCGAGCCCCCGAAGAGGCCATTGCGGAGCC 134
OY	174 GGAGAGGGGACTGCGAGAGACACCCCGCGTCCGGGCTCCGGGTGCCAGGCGCTATGAGGC 233
DG	135 GGAAGGGGAGTGTGAGAGAGACCCCGCGTCCGGGCTCCGGGTGCCAGGCGCTATGAGGC 194
OY	234 CACTCTCTGCTCTGCTCTCTCTGCGGCGCTTGGGGCGCGGCTCGCCCCCACTGAGACACA 293
DG	195 CACTCTCTGCTCTGCTCTCTCTCTGCGGCGCTTGGGGCGCGGCTCGCCCCCACTGAGACACA 254
OY	294 AGATGCCAGCCTTGCCCGGGGACCCCGGCTTCCAGGAGAGCGGGGCGCACATGAGCA 353
DG	255 AGATGCCAGCCTTGCCCGGGGACCCCGGCTTCCAGGAGAGCGGGGCGCACATGAGCA 314
OY	354 GCCAGGCTTCCCGGCGCGGATGGCGCGACAGGCGCGGAGCGCGCGCGGGGCTTCGG 413
DG	315 GCCAGGCTTCCCGGCGCGGATGGCGCGACAGGCGCGGAGCGCGCGGGGCTTCGG 374
OY	414 GAGAGAAAGCGAGGCGGGGAGGCGGAGTGCCTGGGAGACTGAGAGGAGACCCGGGCGCG 473
DG	375 GAGAGAAAGCGAGGCGGGGAGG-CGGAGACTGCCGGGAGCTCAGAGGAGACCCCGGGCGCG 433

OY	774	GAGGAGAGGGGGACCCCGGGGCCACCGGGCCCTGC	GGGAGATGCTCGGTGCCTCCG	533
Dd	434	GAGGAGAGGGGGACCCCGGGGCCACCGGGCCCTGC	GGGAGATGCTCGGTGCCTCCG	493
OY	534	GATCGCCCTTCAAGCGCAGAAGCGCTCCGAGACC	CGGGTGCTCCGCGCTGTACGACACCT	593
Dd	494	GATCGCCCTTCAAGCGCAGAAGCGCTCCGAGACC	CGGGTGCTCCGCGCTGTACGAGCACCT	553
OY	594	TGCCCCCTTCACCGCTGCTGCTGGTAAGCAGCAG	CAGGACATTTACGAGCGCGTCAACGGCAAGT	653
Dd	554	TGCCCCCTTCACCGCTGCTGCTGGTAAGCAGCAG	CAGGACATTTACGAGCGCGTCAACGGCAAGT	613
OY	654	TCACCTGCAGATGGCTGGGCTACTACTTCCGCC	CCCATGGCACCCGCTAACCGGCCA	713
Dd	614	TCACCTGCAGATGGCTGGGCTACTACTTCCGCC	CCCATGGCACCCGCTAACCGGCCA	673
OY	714	GCCTGCACCTTGTATCTGTGTGAAGANTGGCA	AATTCATTCGCTCTTCTTCCAGTTTTCG	773
Dd	674	GCCTGCACCTTGTATCTGTGTGAAGANTGGCA	AATTCATTCGCTCTTCTTCCAGTTTTCG	733
OY	774	GGGGGTGGCCCAACGACCCCTCGCTCCGGGGGG	GGGCCATATGTGAGGCTGGAGCCTTGAG	833
Dd	734	GGGGGTGGCCCAACGACCCCTCGCTCCGGGGGG	GGGCCATATGTGAGGCTGGAGCCTTGAG	793
OY	834	ACCAAAGTGTGGGTGCAGGTGGGTGGGTGAGT	TACATTGGCATTCATCCACGATCAAGA	893
Dd	794	ACCAAAGTGTGGGTGCAGGTGGGTGGGTGAGT	TACATTGGCATTCATCCACGATCAAGA	853
OY	894	CAGACAGACACCTTCTCCGGATTTCTGTGTACT	CCGACGTCCCGCACTCTTGTG	953
Dd	854	CAGACAGACACCTTCTCCGGATTTCTGTGTACT	CCGACGTCCCGCACTCTTGTG	913
OY	954	CTTAGTGGCCACATGCAAAATGACGCTCATGCT	CACTCTAGAAAGGAGGTGTAGGCTG	1013
Dd	914	CTTAGTGGCCACATGCAAAATGACGCTCATGCT	CACTCTAGAAAGGAGGTGTAGGCTG	973
OY	1014	ACAACCAAGTTCATCCAGSAGSGGCTGGCCCC	CTGGAAATATTGTGAATGACTTAGGAGGTG	1073
Dd	974	ACAACCTGTGCATCCAGSAGSGGCTGGCCCC	CTGGAAATATTGTGAATGACTTAGGAGGTG	1033
OY	1074	GGGTAGAGACATCTCCGTCCTGTCTGTGGCA	GAAGTAATGGAGACGTGCTGGATTC	1133
Dd	1034	GGGTAGAGACATCTCCGTCCTGTCTGTGGCA	GAAGTAATGGAGACGTGCTGGATTC	1093
OY	1134	AGGTCTGGCAGCANTGGGCACTGGCTGGATT	TCTGCCCAAGACCAAGAGATGTGCTGTG	1193
Dd	1094	AGGTCTGGCAGCANTGGGCACTGGCTGGATT	TCTGCCCAAGACCAAGAGATGTGCTGTG	1153
OY	1194	CTGGCAAGGTAAAGCCCCCAAGTTGCTCTGTG	TCAGAGGCCACGGTGGGTGCTCTCTT	1253
Dd	1154	CTGGCAAGGTAAAGCCCCCAAGTTGCTCTGTG	TCAGAGGCCACGGTGGGTGCTCTCTT	1213
OY	1254	CCTGGTCTCTGCTTCTGTGATTCCTCCACCC	CCCTCTCTGGTCCGGGGCGGCGCTTT	1313
Dd	1214	CCTGGTCTCTGCTTCTGTGATTCCTCCACCC	CCCTCTCTGGTCCGGGGCGGCGCTTT	1273
OY	1314	TCTCAGAGATACATCAATAAAGTAAGAACCT	CATTAATAAAAAAAAAAAAAAAA	1368
Dd	1274	TCTCAGAGATACATCAATAAAGTAAGAACCT	CATTAATAAAAAAAAAAAAAAAA	1328
 RESULT 12 AAF45000 standard; cDNA; 1338 BP. XX AAF45000; XX DT 28-MAR-2001 (first entry) XX DE Human secreted protein related coding sequence SEQ ID NO: 107. XX KW Human; mouse; secreted protein; TANGO25; TANGO 257; TANGO 281;				

100

D	b		494	GATTCGCCCTTGACGCGCCAAAGCSCCTCCGAAGGCCGGGTGCTCTCCGCCGTCTGAGGCACCCT	553
O	y		594	TGCCTTTGACCCGCGTGTGTGAACGAGACGAGCAATATTAGACGCCGTACCGGCAAGT	653
D	b		554	TGCCCTTTCACCGCGCTGTGCGAAMCGACGAGGACATTTACGACGCCGTACACCGCAAGT	613
O	y		654	TCACCTGCCAGGTGGCTGGGCTACTACTCTTCGCGCTTCATCCGACCCTGTACCGGGCCA	713
D	b		614	TCACCTGCCAGGTGGCTGGGCTACTACTCTTCGCGCTTCATCCGACCCTGTACCGGGCCA	673
O	y		.714	GCCTCAGTTTGATCTGGTGTGAAGAATGGCGAATATCATTTGCTCTTTCTTCAGATTTTTCG	773
D	b		674	GCCTCAGTTTGATCTGGTGTGAAGAATGGCGAATATCATTTGCTCTTTCTTCAGATTTTTCG	733
O	y		774	GGGGGTGGCCCCAACGCGACCCCTGCTTCGCGGGGGGGCCATGATGAGCTGGAGCCCTGAGC	833
D	b		734	GGGGGTGGCCCCAACGCGACCCCTGCTTCGCGGGGGGGCCATGATGAGGCTGGAGCCCTGAGC	793
O	y		834	ACCAAATGTGGGTGTCAGATGGGTGTGGGTGACTACATAATGGCACTCTATTGGCACATTCACA	893
D	b		794	ACCAAATGTGGGTGTCAGATGGGTGTGGGTGACTACATAATGGCACTCTATTGGCACATTCAGA	853
O	y		894	CAGACACACACCTTCTCCGGAATTTCTGCTGTACTCCGACATGGCACAAGCTCCCGACGTTTTG	953
D	b		854	CAGACACACACCTTCTCCGGAATTTCTGCTGTACTCCGACATGGCACAAGCTCCCGACGTTTTG	913
O	y		954	CTTAGTCCCGACATGCAAAAGTAGAGCTATGCTGCATCTCTTAAGAGAGAGGTGTAGGCTG	1013
D	b		914	CTTAGTCCCGACATGCAAAAGTAGAGCTATGCTGCATCTCTTAAGAGAGAGGTGTAGGCTG	973
O	y		1014	ACAACGAGGTCATCCAGSAGAGGGCTGGCCCCCTCGCAATATTGTGTGAATGACTAGGAGAGTG	1073
D	b		974	ACAACCTGGTCATCCAGSAGAGGGCTGGCCCCCTCGCAATATTGTGTGAATGACTAGGAGAGTG	1033
O	y		1074	GGGTAAAGCACTCTCCGCTCTGCTGCTGCGCAAGAATGGGAACAAGTGGCTGTGCGCATC	1133
D	b		1034	GGGTAAAGCACTCTCCGCTCTGCTGCTGCGCAAGAATGGGAACAAGTGGCTGTGCGCATC	1093
O	y		1134	AGGTCTGCGAGCATGGGGGAGTGGCTGTGATTTTGCCCAACAGACAGAGAGATGTGCTGTG	1193
D	b		1094	AGGTCTGCGAGCATGGGGGAGTGGCTGTGATTTTGCCCAACAGACAGAGAGATGTGCTGTG	1153
O	y		1194	CTGGCAAGTGTAAGTCCCCCAGTTGCTGTGCTCAGAGGCCACAGCGTGGGTCTCTCTT	1253
D	b		1154	CTGGCAAGTGTAAGTCCCCCAGTTGCTGTGCTCAGAGGCCACAGCGTGGGTCTCTCTT	1213
O	y		1254	CCTGGTCTCTGCTTCTCTCTGATCTCTGCCACCCCGCTCTGCTCTCGTGCGGCGCGCCCTTT	1313
D	b		1214	CCTGGTCTCTGCTTCTCTCTGATCTCTGCCACCCCGCTCTGCTCTCGTGCGGCGCGCCCTTT	1273
O	y		1314	TCTCAGAGATCACTCAATTAACCTAAGAGAACCTCTCATTAACAAAAAAAAAAAAAAA	1368
D	b		1274	TCTCAGAGATCACTCAATTAACCTAAGAGAACCTCTCAAAAAAAAAAAAAAAA	1328
RESULT 13					
AAFA45001					
ID	AAF45001 standard; cDNA; 1338 BP.				
XX	AAF45001;				
XX	28-MAR-2001 (first entry)				
XX	Human secreted protein related coding sequence SEQ ID NO: 109.				
KM	Human; mouse; secreted protein; TANGO253; TANGO 257; TANGO 281;				
KW	INTERCEPT 258; coronary disorder; olfactory disorder;				
KX	neurological disorder; pulmonary disorder; immunological disorder;				
XX	developmental disorder; kidney disorder; ss.				
OS	Homo sapiens.				
N	MO200078808-A1.				

SV Sequence 1338 BP; 228 A; 424 C; 460 G; 226 T; 0 other;

5A CCCCCCCTTCCCGGGG.....
0V

[illegible][illegible]

RESULT 14	
AAL39624	
ID	AAL39624 standard; DNA; 3248 BP.
XX	
AC	AAL39624;
XX	
D7	05-SEP-2002 (first entry)
XX	
DE	Human secreted protein DNA SEQ ID NO 68.
XX	
KW	Antiarteriosclerotic; cytosstatic; HIV; antiallergic; antinaeemic;
KW	antisthmatic; cardiant; vasotropic; neuroprotective; nootropic; SSCP;
KW	anticonvulsant; antiparkinsonian; cerebroprotective; antiinflammatory;
KW	immunosuppressive; human secreted protein; cell proliferative disorder;
KW	arteriosclerosis; cancer; autoimmune; inflammatory disorder; AIDS;
KW	allergy; anaemia; asthma; cardiovascular disease; developmental disorder;
KW	ischemic heart disease; congestive heart failure; neurological disorder;
KW	renal tubular acidosis; hypothyroidism; Alzheimer's disease; dementia;
KW	Parkinson's disease; epilepsy; stroke; knockin humanised animal;
KW	transgenic animal; gene therapy; gene; ds.
XX	
OS	Homo sapiens.
XX	
XX	
PN	WO200238602-A2.
XX	

PD 16-MAY-2002.
 XX 08-NOV-2001; 2001WO-US47420.
 XX 08-NOV-2000; 2000US-247505P.
 PR 09-NOV-2000; 2000US-248642P.
 PR 16-NOV-2000; 2000US-249824P.
 PR 21-NOV-2000; 2000US-252824P.
 PR 08-DEC-2000; 2000US-254305P.
 PR 18-DEC-2000; 2000US-256448P.
 PA (INCYTE) INCYTE GENOMICS INC.
 XX Yue H, Yao MG, Gandhi AR, Baughn MR, Swarnakar A, Walla NK;
 PI Sanjanwala M, Thornton M, Elliott VS, Lu Y, Gietzen KJ, Burford N;
 PI Ding L, Hefalla AJA, Yang YT, Bandman O, Warren BA, Honchell CD;
 PI Lu DM, Thangavelu K, Lee S, Xu Y, Yang J, Lal PG, Tran B;
 PI Ison CH, Duggan BM, Saperstein SK;
 DR MPI: 2002-519296/55.
 DR P-PSDB; AAO21663.
 XX Human secreted proteins and polynucleotides for diagnosing, treating or
 PT preventing disorders of cell proliferative, cardiovascular,
 PT developmental, neurological and autoimmune/inflammatory disorders -
 PS Claim 5; Page 194-195; 229pp; English.
 CC The invention relates to an isolated human secreted protein (SECP)
 CC polypeptide from 63 fully defined protein sequences given in the
 CC specification. The polypeptide is useful for the diagnosing/treating of a
 CC disease with decreased/overexpression of SECP. Examples of disorders
 CC associated with abnormal expression of SECP include a cell proliferative
 CC disorder e.g. arteriosclerosis, cancer; autoimmune/inflammatory
 CC congestive heart failure, ischemic heart disease; developmental disorder
 CC e.g. renal tubular acidosis, hypothyroidism; neurological disorder e.g.
 CC Alzheimer's disease, dementia, Parkinson's disease, epilepsy or stroke.
 CC The SECP polynucleotide and polypeptide are further useful for analysing
 CC the proteome of a tissue or a cell type. The polynucleotide is useful for
 CC creating knockin humanized animals (pigs) or transgenic animals (mice or
 CC rats) to model human disease, and for somatic or germline gene therapy,
 CC and further for generating hybridisation probes useful in mapping the
 CC naturally occurring genomic sequence. This polynucleotide sequence
 CC represents the DNA of a human secreted protein of the invention.
 CC
 CC Sequence 3248 BP; 565 A; 1058 C; 946 G; 679 T; 0 other:
 SO
 Query Match 78.8%; Score 1084.8; DB 24; Length 3248;
 Best Local Similarity 90.5%; Pred. No. 1.1e-190;
 Matches 1222; Conservative 0; Mismatches 2; Indels 126; Gaps 1;

Db	2325		
Qy	367	GGGCGGATGGCCCGGACGCGCCGACGGCGCCCGGGGGCTCCGGAGAGAAAGCGCA	2328
Db	2329	-----	2328
Qy	427	GGGCGGAGCGCGGACATGCGGAGCTCGAGGAGAACCCCGGGCGCGGAGAGAGCGGG	486
Db	2329	-----CCGGAGCTCGGGGACCTCGAGGGAGACCCCGGGCGCGGAGAGAGCGGG	2378
Qy	487	ACCGCGGGGGCCACCGGGGCTCCCGGAGTGTCTGCTCCGCGATCCGCTTCCAG	546
Db	2379	ACCGGGGGGGCCACCGGGGCTCCCGGAGTGTCTGCTCCGCGATCCGCTTCCAG	2438
Qy	547	CGGCAAGCGCTCCGAGAGCGGGGCTCCGCGCTGTGACGACACCTTGGCTTGCACG	606
Db	2439	CGGCAAGCGCTCCGAGAGCGGGGCTCCGCGCTGTGACGACACCTTGGCTTGCACG	2498
Qy	607	CGTGTGTGAAG	666
Db	2499	CGTGTGTGAAG	2558
Qy	667	GGCTGGGCTCTACTACTTCCGCGCTCCATGCCACCTTACCGGGCCAGCTTGA	726
Db	2559	GGCTGGGCTCTACTACTTCCGCGCTCCATGCCACCTTACCGGGCCAGCTTGA	2618
Qy	727	TCTGTGAAG	786
Db	2619	TCTGTGAAG	2678
Qy	787	GGCAGCTGCTGCTGG	846
Db	2679	GGCAGCTGCTGCTGG	2738
Qy	847	GCAGTGGTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	906
Db	2739	GCAGTGGTGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2798
Qy	907	CTCCGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	966
Db	2799	CTCCGATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2858
Qy	967	GCAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1026
Db	2859	GCAAGTGAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	2918
Qy	1027	CCAGAGAGGG	1086
Db	2919	CCAGAGAGGG	2978
Qy	1087	TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1146
Db	2979	TCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3038
Qy	1147	TGGGCGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1206
Db	3039	TGGGCGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3098
Qy	1207	GTCGCCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1266
Db	3099	GTCGCCAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3158
Qy	1267	TTCCTGATCTCTCCGACCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1326
Db	3159	TTCCTGATCTCTCCGACCCCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	3218
Qy	1327	TCAATTAACCTAAGAACCTTCAAAAAA 1356	
Db	3219	TCAATTAACCTAAGAACCTTCAAAAAA 3248	

RESULT 15

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2003, 03:22:25 ; Search time 1901 Seconds
(without alignments)
11731.292 Million cell updates/sec

Title: US-09-943-851A-41

Perfect score: 1377
Sequence: 1 gacacgtctctcttgagctct.....aaaaaaaaaaaaaaaa 1377

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 33308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1:	em_estbda:*
2:	em_esthum:*
3:	em_estlin:*
4:	em_estlmu:*
5:	em_estloy:*
6:	em_estlpl:*
7:	em_estlro:*
8:	em_estlro:*
9:	em_estlro:*
10:	em_estlro:*
11:	em_estlro:*
12:	em_estlro:*
13:	em_estlro:*
14:	em_estlro:*
15:	em_estlro:*
16:	em_estlro:*
17:	em_estlro:*
18:	em_estlro:*
19:	em_estlro:*
20:	em_estlro:*
21:	em_estlro:*
22:	em_estlro:*
23:	em_estlro:*
24:	em_estlro:*
25:	em_estlro:*
26:	em_estlro:*
27:	em_estlro:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	767.8	55.8	869	13	B1763193
2	750.6	54.5	1154	13	BMS47549
3	732.8	53.2	1093	14	BM924569
4	730.2	53.0	846	13	B1458455
5	712.6	51.8	801	13	B1490880
6	672.2	48.8	724	13	B1771879

7	660.4	48.0	1045	14	BM920874
8	641.8	46.6	792	13	B1770921
9	631.1	45.8	887	13	B1490062
10	607.2	44.1	1623	14	BM926477
11	596.4	43.3	596	14	BM926609
12	586.4	42.6	596	14	BM926609
13	583.2	42.4	941	13	BM926609
14	578.2	42.0	578	14	BM926609
15	577.2	41.9	578	14	BM926609
16	572.4	41.6	574	14	BM926609
17	572.4	41.6	574	14	BM926609
18	554	40.2	870	9	AF451167
19	551.4	40.0	553	12	BF882978
20	532	38.6	532	12	BF724241
21	521	37.5	521	14	BM926609
22	517	37.5	517	14	BM926609
23	507.6	36.9	574	10	AM150070
24	507.6	36.9	574	10	AM150070
25	489.4	35.5	567	13	BM926609
26	479	34.8	778	12	BF098614
27	463.4	33.7	867	13	B1818537
28	457.4	33.2	1349	14	B0673882
29	453.2	32.9	702	12	BF055285
30	452.4	32.9	454	9	AT805087
31	448.4	32.6	462	10	AM070344
32	440.2	32.0	445	14	BM673416
33	438.4	31.8	842	13	B1668017
34	426.4	31.0	449	13	BM668611
35	424.4	30.8	442	14	BM668611
36	423.4	30.7	548	12	BF906443
37	421.6	30.6	630	12	BF906443
38	409.8	29.8	413	9	AT083823
39	408.4	29.7	410	14	BM707291
40	405.4	29.4	408	9	AT083824
41	402.4	29.2	405	9	AT083824
42	395.6	28.7	547	12	BE722922
43	394.6	28.7	936	12	BF540263
44	392	28.5	392	10	BE18819
45	387.2	28.1	392	10	AM192716

ALIGNMENTS

RESULT 1
LOCUS B1763193 869 bp mRNA linear EST 25-SEP-2001
DEFINITION 603049929F1 NIH_MGC_116 Homo sapiens CDNA clone IMAGE:5190222 5',
ACCESSION B1763193
VERSION B1763193.1 GI:15754771
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE 1 (bases 1 to 869)
AUTHORS NIH-MGC http://mgc.ncl.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@ncl.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1475 row: 1 column: 07
High quality sequence stop: 777.
Location/Qualifiers
I. 869

FEATURES
source

Db 661 CCTAGAAAGAGGGGTTGTAGGCTGACCAACCGGTCATCCAGAGGGGCTGGGCCCTCGA 720
 QY 1050 ATTTTGTGAATGACTAGAGGAGTGGGGTAGAGCACTCTCCGCTCGTGGCTGGCAAGAA 1100
 Db 721 ATTTTGTGAATGACTAGAGGAGGGGGGTANMANCCTCTCCGTCCTGGTGGCA 780
 QY 1110 TG 1111
 Db 781 GG 782.

RESULT	4
B1458455	
LOCUS	B1458455
DEFINITION	6031988377r1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:5276184 5' ,
ACCESSION	B1458455
VERSION	B1458455.1 GI:15249111
KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

1 (bases 1 to 846)
NII-MGC <http://mgc.ncl.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg Ph.D.

Source: Mikiros Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NIGRI), Shirok
Toshiyuki and Piero Carninci (IKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
Plate: LCLM1702 row: 1 column: 09
High quality sequence strip: 815.

```

FEATURES
SOURCE
Location/Qualifiers
1. 846
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone IMAGE:5278184"
/clone_11b="NH_MGC_96"
/clone_type="Hypothalamus"
/lab_host="DH10B"
/notes="Organ: brain; Vector: pBluescript (modified
pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtgaga
5'); Oligo-dT primed using primer 5' -gtgtgtgtgtgtgtgttvtvn-3',
size-selected for average insert size 2.3 kb and
normalized to R07 5. This is a primary library enriched
for full-length clones and constructed using the
Cap-trapper method (Carninci, in preparation). Library
constructed by M. Brownstein (NIH/NHGRI, National
Institutes of Health). Note: this is a NH_MGC Library."
129 a 271 c 330 g 115 t 1 others
BASE COUNT
GCIN

```

Query Match	53.0%	Score 730.2	DB 13	Length 846
Best Local Similarity	96.1%	Pred. No. 1.1e-90		
Matches 803	Conservative 0	Mismatches 24	Indels 9	Gaps 5
QY	7	TTCTCTTGAGTCTGGAGAGAAACGGACGGCGAGGAGACCGAAACCGAGACTGCGGT	66	
DB	19	TCCTCTTGAGTCTGGAGAGAAACGGACCGCGAGGAGACCGAAACCGAGACTGCGGT	78	
QY	67	GACGGCAGGCGCAGGGGGCGCCCTGGCCGGGAGAAAGCGCGGGGCTTGAGCACCACCAACT	126	
DB	79	GACGGCAGGCGCAGGGGGCGCCCTGGCCGGGAGAAAGCGCGGGGCTTGAGCACCACCAACT	138	
QY	127	GGAGGCTCCGAGTACGAGAGCGCCCGCGAAGAGACCCCATGGGGAGACCGGGAGGGGGACT	186	

[illegible]

```

RESULT 5
BI430880/c
LOCUS
DEFINITION
BI430880 801 bp mRNA
G030186771 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:512880 3',
ACCESSION
BI430880
VERSION
BI430880.1 GI:15330108
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 801)
AUTHORS
NIH-MGC http://mgc.ncl.nih.gov/.
JOURNAL
National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
cDNA distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov

```

Plate: LLAM11430 row: f column: 17
High quality sequence start: 8
High quality sequence stop: 738.

BASE COUNT	ORIGIN
180 a	251 c
237 g	133 t

NOTE: THIS IS A NIH/MGC Library.

Query Match	51.8%;	Score 712.6;	DB 13;	Length 801;
Best Local Similarity	98.28;	Pred. No. 2.7e-88;		
Matches 773; Conservative	0;	Mismatches 0;		

Db	127	AGTGTGCTGTGCTGGCAAGTGTAAAGTCCGCCACAGTGTCTGTGTCCAGAGAACCCACAGTGG	68
QY	1243	GGTCTCTCTTTCCTGATCTCTGCTCTGTGATTCCTCCACCCACCCCTCTGCTCTGG	13
Db	67	GGTCTCTCTTTCCTGCTCTCTGCTCTGTGATTCCTCAATAACCCCTCTGCTCTGG	8
QY	1303	GCCGCGCC	1309
Db	7	GGGGGCC	1

RESULT 6	BI771879	724 bp	mRNA	linear	EST-25-SEP-2001
LOCUS	BI771879				
DEFINITION	BI771879				
	603055280F1 NIH_MGC_122		Homo sapiens	cDNA clone IMAGE:5204971	5'
ACCESSION	BI771879				
VERSION	BI771879.1				
KEYWORDS	EST.				
SOURCE	human.				
ORGANISM	Homo sapiens				

FEATURES
SOURCE

BASE COUNT	clnt IS d	NIH-MGC Library	
176 a	184 c	211 g	152 t
ORIGIN			1 others

Query Match	48.8%;	Score 672.2;	DB 13;	Length 724;
Best Local Similarity	99.4%;	Pred. No. 9.2e-83;		
Matches 674;	Conservative	0;	Mismatches 4;	Indels 0;

QY 880 TCCAGCATCAAGACAGACAGACACTTCTCCGATTTCTGTGACTCCGACATGGACAG 939
 DB 182 TCCAGCATCAAGACAGACAGACACTTCTCCGATTTCTGTGACTCCGACATGGACAG 241
 QY 940 CTCCTCCAGTCTTCTGCTTGTAGTCCGACATGCAAAAGTCAAGTCTCTCACTCTCTAGAG 999
 DB 242 CTCCTCCAGTCTTCTGCTTGTAGTCCGACATGCAAAAGTCAAGTCTCTCACTCTCTAGAG 301
 QY 1000 AGGCTGTGAGCTGACAGACAGTCTATCCAGAGAGGCTGCCCCCTGGAAATATGTA 1059
 DB 302 AGGCTGTGAGCTGACAGACAGTCTATCCAGAGAGGCTGCCCCCTGGAAATATGTA 361
 QY 1060 TGACTAGGAGGAGGAGGAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1119
 DB 362 TGACTAGGAGGAGGAGGAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 421
 QY 1120 GCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1179
 DB 422 GCTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 481
 QY 1180 AGGAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 1239
 DB 482 AGGAGTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCT 541
 QY 1240 TGGGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 1299
 DB 542 TGGGCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCT 601
 QY 1300 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1359
 DB 602 GGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 661
 QY 1360 AAAAAAAAAAAAAAAAAA 1377
 DB 662 AAAAAAAAAAAAAAAAAA 679

RESULT 7
 BM920874
 LOCUS 1045 bp mRNA linear EST 12-MAR-2002
 DEFINITION AGENCOURT.6706053 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5752141
 ACCESSION BM920874
 VERSION 5', mRNA sequence.
 KEYWORDS BM920874.1 GI:19371253
 SOURCE EST.
 ORGANISM human.
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
 NIH-MGC http://mgc.ncl.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cga@bhs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL12785 row: n column: 14
 High quality sequence stop: 415.
 Location/Qualifiers
 1. 1045
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="5752141"
 /clone_lib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector:"

pcmw-sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH-MGC Library."

BASE COUNT 176 a 343 c 341 g 183 t 2 others
 ORIGIN

Query Match 48.7%; Score 671; DB 14; Length 1045;
 Best Local Similarity 95.4%; Pred. No. 1.1e-82;
 Matches 733; Conservative 0; Mismatches 31; Indels 4; Gaps 4;

QY 171 GCCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 230
 DB 37 GCCTGTCCAGACAGACTGCCAGAGAGACCCGCGGCTCCGCTCCGCTCCGCTCCGCTATGA 96
 QY 231 GCCACTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 290
 DB 97 GGCCTACTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 156
 QY 291 ACAAGATCCCAAGCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 350
 DB 157 ACAAGATCCCAAGCCCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 216
 QY 351 GGAGCAGAGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 410
 DB 217 GCAGCAGAGGCTTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 276
 QY 411 CGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 470
 DB 277 CGGAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 336
 QY 471 CGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 530
 DB 337 CGCAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 396
 QY 531 CGCAGATCCGCTTACAGCCCAAGCGCTCCGAGAGCCGCGGCTCCGCTCCGCTCCGCTCCGCTCC 590
 DB 397 CGCAGATCCGCTTACAGCGCCCAAGCGCTCCGAGAGCCGCGGCTCCGCTCCGCTCCGCTCCGCT 456
 QY 591 CTTGCGCTTGCAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 650
 DB 457 CTTGCGCTTGCAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 516
 QY 651 AGTTCACTTGCAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 710
 DB 517 AGATCACCTTGCAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 576
 QY 711 CGACCTCGAGTTGATCTGCGCGAAGAGGCGCAATTCATTCGCTCTGCTCTGCTCTGCTCTGCT 770
 DB 577 CGACCTCGAGTTGATCTGCGCGAAGAGGCGCAATTCATTCGCTCTGCTCTGCTCTGCTCTGCT 636
 QY 771 TCGGGGCGTGGCCCAAGCCAGCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 830
 DB 637 TCGGTGGTGGCCCAAGCCAGCTCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 696
 QY 831 AGAGCCAGTGTGGTGGAGGTG-GGTGTGGGTGACTACA-TTGGCATCTATGCCA-GCA 887
 DB 697 ATGACCAAGTGTGGTGGAGGTG-TGGTGTGGGTGACTACA-TTGGCATCTATGCCA-TTCA 756
 QY 888 TCAAGACAGACAGACACTTCT-CCGATTTCTGGTGTGCTCCGATGG 934
 DB 757 TCCAGACAGACAGACACTTCTCCCGATTTCTGGTGTGACTCCGCTGG 804

RESULT 8
 LOCUS B1770921
 DEFINITION 603059737F1 NIH_MGC_122 Homo sapiens cDNA clone IMAGE:5209065 5',
 B1770921 792 bp mRNA linear EST 25-SEP-2001

mRNA sequence.
 BI770921
 VERSION BI770921.1 GI:15762499
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 NIH (bases 1 to 792)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11524 row: 1 column: 10
 High quality sequence stop: 782.
 Location/Qualifiers
 1..792
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:520905"
 /clone_lib="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-SPORT6;
 Site_1: NotI; Site_2: EcoRV (destroyed); RNA source:
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invtrogen). Research Genetics tracking code 026. Note:
 this is a NIH-MGC Library."

BASE COUNT 142 a 232 c 237 g 180 t 1 others

ORIGIN

Query Match 48.0%; Score 660.4; DB 13; Length 792;
 Best Local Similarity 96.9%; Pred. No. 3.5e-81;
 Matches 758; Conservativity 0; Mismatches 16; Indels 8; Gaps 8;

QY 562 GACCGGAGGCGCCGCGCTGAGCAGCCCTTGGCTTCGACCCGCTGCTGTAACGA 621
 |||||||
 DB 1 GACCGGAGGCGCTGCGCCGCTGAGCAGCCCTTGGCTTCGACCCGCTGCTGTAACGA 60
 |||||||

QY 622 GCAGGAGCATTAAGAGCGCGTACCGGCAAGTTCACTGCGCGAGTGGCTGGGCTACTA 681
 |||||||
 DB 61 GCAGGAGCATTAAGAGCGCGTACCGGCAAGTTCACTGCGCGAGTGGCTGGGCTACTA 120
 |||||||

QY 682 CTTCCGCGGTCATGCGACCGTGTACCGGGCCAGCTGCAAGTTGATGCTGTAAGAATGG 741
 |||||||
 DB 121 CTTCCGCGGTCATGCGACCGTGTACCGGGCCAGCTGCAAGTTGATGCTGTAAGAATGG 180
 |||||||

QY 742 CGAATCATTCGCTCTTCTTCGATTTTGGGGGGGGGGCCAAAGCAGCCTGCTCTC 801
 |||||||
 DB 181 CGAATCATTCGCTCTTCTTCGATTTTGGGGGGGGGGCCAAAGCAGCCTGCTCTC 240
 |||||||

QY 802 GGGGGGGGCATGTGAGCGTGAAGCCTGAGAGCAAGTGTGGGTGAGAGTGGGTGGG 861
 |||||||
 DB 241 GGGGGGGGCATGTGAGCGTGAAGCCTGAGAGCAAGTGTGGGTGAGAGTGGGTGGG 300
 |||||||

QY 862 TGACTACATTTGGCATCTATGCGACATCAAGACAGACAGACCTTCTCCGATTTCTGT 921
 |||||||
 DB 301 TGACTACATTTGGCATCTATGCGACATCAAGACAGACAGACCTTCTCCGATTTCTGT 360
 |||||||

QY 922 GTACTCCAGCTGGAGACAGCTCCCAAGTCTTGTAGTGGCACTGCAAAAGTGGAGCAT 981
 |||||||

DB 361 GTACTCCAGCTGGACAGCTCCCTCTTCTTGTAGTCCCACTGCAAAAGTAGCTCAT 420
 |||||||

QY 982 GCTCTACTCTCTTAAGAGAGGTGTGAGCTGACACACAGATCATCAGAGAGGC-TGGC 1040
 |||||||

DB 421 GCTCTACTCTCTTAAGAGAGGTGTGAGCTGACACACAGATCATCAGAGAGGC-TGGC 480
 |||||||

QY 1041 CCCCCTGGATATTGTAATGACTAGAGAGGTGGGTAGAGC-ACTCTCCGCTGCTGTC 1099
 |||||||

DB 481 CCCCCTGGATATTGTAATGACTAGAGAGGTGGGTAGAGC-ACTCTCCGCTGCTGTC 540
 |||||||

QY 1100 TGGCAAGATGGAAGACAGTGGCTGCTGTCGATCAGAGGTGTG-CAGCATGGGAGCTGGC 1158
 |||||||

DB 541 TGGCAAGATGGAAGACAGTGGCTGCTGTCGATCAGAGGTGTG-CAGCATGGGAGCTGGC 600
 |||||||

QY 1159 TGGATTTTCCGCAAGACAGCAG-AGAGAGTGTGCTGCTGTCGCAAGTG-TAAGTCCCAAGT 1216
 |||||||

DB 601 TGGATTTTCCGCAAGACAGCAG-AGAGAGTGTGCTGCTGTCGCAAGTG-TAAGTCCCAAGT 660
 |||||||

QY 1217 T-GCTCTGTCGAGAGAGCCACAG-GTGGGCTGCTCTT-CCTGCTCTGCTGCTG 1273
 |||||||

DB 661 TAGCTCTGTCGAGAGAGCCACAGAGTGTGCTGCTTACCTGCTGCTGCTGCTGCTG 720
 |||||||

QY 1274 GATCTCTCCCAACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1333
 |||||||

DB 721 GATCTCTCCCAACCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
 |||||||

QY 1334 AC 1335
 ||

DB 781 AC 782

RESULT 9
 BI490062 887 bp mRNA linear EST 28-AUG-2001
 BI490062 60303186761 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5172880 5',
 LOCUS mRNA sequence.
 DEFINITION BI490062
 mRNA sequence.
 ACCESSION BI490062.1 GI:15329290
 VERSION EST.
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 NIH (bases 1 to 887)
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: LLM11430 row: 1 column: 17
 High quality sequence stop: 815.
 Location/Qualifiers
 1..887
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5172880"
 /clone_lib="NIH_MGC_115"
 /lab_host="DH10B"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and
 enriched for full-length clones and was constructed by C.

FEATURES
 source

Gruber (Invitrogen). Research Genetics tracking code
 BASE COUNT 146 a 277 c 307 g 157 t
 ORIGIN

Query Match 46.68; Score 641.8; DB 13; Length 887;
 Best Local Similarity 84.78; Pred. No. 1.1e-78;
 Matches 858; Conservative 0; Mismatches 22; Indels 133; Gaps 7;

```

OY 55 CAGGACTGGGTGACGCGCAGGAGGCGGCGCTGCGCGGAGAAAGCCGGGGCTGGA 114
DB 1 CAGGACTGGGTGACGCGCAGGAGGCGGCGCTGCGCGGAGAAAGCCGGGGCTGGA 60
OY 115 GCACCACTGAGAGGCTCCGAGTAGAGCGCCCGCAAGAGAGGAGCATCGGGAGCGG 174
DB 61 GCACCACTGAGAGGCTCCGAGTAGAGCGCCCGCAAGAGAGGAGCATCGGGAGCGG 120
OY 175 GGAGGGGGGACTGCGAGAGAGAGCCCGGCGCTGCGCGGAGAAAGCCGGGGCTGGA 234
DB 121 GGAGGGGGGACTGCGAGAGAGAGCCCGGCGCTGCGCGGAGAAAGCCGGGGCTGGA 180
OY 235 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 294
DB 181 ACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
OY 295 GATCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 354
DB 241 GATCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 256
OY 355 CGAGGGCTTGGGGGGGGGGGATGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGG 414
DB 257 ----- 256
OY 415 AGAGAAAGCGGAGCGCGGAGCGCGGAGCTGCGCGGAGCTGAGAGGAGCCCGGCGCG 474
DB 257 ----- 294
OY 475 AGGAGAGCGGAGCGCGGAGCGCGGAGCGCGGAGCTGAGAGGAGCGCGCGCGCGG 534
DB 295 AGGAGAGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCTGAGAGGAGCGCGCG 354
OY 535 ATCCGCGCTTACAGCGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGG 594
DB 335 ATCCGCGCTTACAGCGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGGAGCGCGG 414
OY 595 GCCCTTGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 654
DB 415 GCCCTTGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 474
OY 655 CACCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
DB 475 CACCTGCGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 534
OY 715 CTTGCACTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 774
DB 535 CTTGCACTTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 594
OY 775 GGGGTGGCCCAACCGGCGCTGCTGCGGGGGGGCGCATGTGAGCTGAGAGCTGAGGA 834
DB 595 GGGGTGGCCCAACCGGCGCTGCTGCGGGGGGGCGCATGTGAGAGCTGAGAGCTGAGGA 653
OY 835 CCAAGTGTGGGTGAGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
DB 654 CCAAGTGTGGGTGAGAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 713
OY 895 AGACAGACACTTCTCGGACTTTC-TGTGTACTCGCACT-GGACAGAGTCCCACTCTTT 952
DB 714 AGACAGACACTTCTCGGACTTTC-TGTGTACTCGCACT-GGACAGAGTCCCACTCTTT 773
OY 953 GCTAGTGGCCCACTGCAAGT--GAGCTCATGCTCTC-ACCTCAGAGAGAGGGTGTGAG 1009
DB 774 GCTAGTGGCCCACTGCAAGT--GAGCTCATGCTCTC-ACCTCAGAGAGAGGGTGTGAG 833

```

OY 1010 GCTGCAACACGAG--TCATCCAGAGAGCGCTGCGCCCTCGTAATATGTGATG 1061
 DB 834 GCTGCAACACAGGTTTCATCCAGAGAGCGCTGCGCCCTCGTAATATGTGATG 886

RESULT 10
 BM926477 1623 bp mRNA linear EST 12-MAR-2002
 LOCUS AGENCOURT_6644538 NIH_MGC_122 Homo sapiens CDNA clone IMAGE:5766864
 DEFINITION 5' mRNA sequence.
 ACCESSION BM926477
 VERSION BM926477.1 GI:19376856
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 1623)
 NIH-MGC <http://mgi.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 found through the I.M.A.G.E. Consortium/LNL at:
<http://image.llnl.gov>
 Plate: LLM12824 row: d column: 01

FEATURES
 source
 Location/Qualifiers
 1..1623
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5766864"
 /clone_lib="NIH_MGC_122"
 /lab_host="DH10B"
 /note="Organ: pooled lung and spleen; Vector: pCMV-Sport6;
 Site_1: NotI; Site_2: EcoRV (destroyed); RNA source
 anonymous pool of 24 week female lung, 16 week female
 spleen, and 20-22 week male spleens. Library is oligo-dT
 primed and directionally cloned (EcoRV site is destroyed
 upon cloning). Average insert size 1.4 kb, insert size
 range 1-3 kb. Library is normalized and enriched for
 full-length clones and was constructed by C. Gruber
 (Invitrogen). Research Genetics tracking code 026. Note:
 this is a NIH-MGC Library."

BASE COUNT 263 a 505 c 550 g 299 t 6 others
 ORIGIN

Query Match 45.88; Score 631; DB 14; Length 1623;
 Best Local Similarity 100.0%; Pred. No. 2.4e-77;
 Matches 631; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 71 GCAGGCGAGGGGGCGCTGCGCGGAGAGAGCGGCGGCTGAGACACCAACTGAG 130
DB 1 GCAGGCGAGGGGGCGCTGCGCGGAGAGAGCGGCGGCTGAGACACCAACTGAG 60
OY 131 GGTCCGAGTACGAGAGCGCCCGAAGAGAGCGGCGGAGCGGAGCGGAGCGGAGCTG 190
DB 61 GGTCCGAGTACGAGAGCGCCCGAAGAGAGCGGCGGAGCGGAGCGGAGCGGAGCTG 120
OY 191 GAGAGCCCGGGGCTCGGCTCCGCTGCGAGCGGTATGAGGCCACTGCTGCTGCTG 250
DB 121 GAGAGCCCGGGGCTCGGCTCCGCTGCGAGCGGTATGAGGCCACTGCTGCTGCTG 180
OY 251 CTCCTGGGCTCGGCGCGCGGCTGCGCCCACTGAGAGCAACAAGATCCCACTGTC 310
DB 181 CTCCTGGGCTCGGCGCGCGGCTGCGCCCACTGAGAGCAACAAGATCCCACTGTC 240
OY 311 CCGGGGACCCCGCGCTTCAGAGCAGCGGGGCCACACATGGACAGGAGGCTTCCGGCG 370

```



```

Db      241  CCGGGGACCCCGGCTTCAGGACCGCGGCGCACATGCGAGGGCTTCGGGCG
QY      371  CCGGATGCGCGGCGACGCGCGCGGCGCGGCGGCTCCGAGAGAGAGAGGCGGCG
Db      301  CCGGATGCGCGGCGACGCGCGCGGCGCGGCGGCTCCGAGAGAGAGAGGCGGCG
QY      431  CCGGAGCGCGGAGCTCCGCGGACCTCGAGGAGAGAGAGAGAGAGAGAGAGAG
Db      361  GGGAGCGCGGAGCTCCGCGGAGCTCGAGGAGAGAGAGAGAGAGAGAGAGAG
QY      491  GCGGGGCGCGGCGGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      421  GCGGGGCGCGGCGGCGGCGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      551  AAGCGCTCCGAGAGCGGCGGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCG
Db      481  AAGCGCTCCGAGAGCGGCGGCTTCGCGGCGGCGGCGGCGGCGGCGGCGGCG
QY      611  CTGGTGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      541  CTGGTGAACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      671  GGGGCTCTACTTCTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG
Db      601  GGGGCTCTACTTCTGCGCGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGG

```

```

RESULT 11
LOCUS    BC706609          630 bp      mRNA      linear      EST 07-MAY-2001
DEFINITION
602672994.F1 NIH_MGC_96 Homo sapiens cDNA clone IMAGE:479558 5',
ACCESSION BC706609
VERSION    BC706609.1 GI:13982121
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
            NIH_MGC http://mgc.ncl.nih.gov/.
            National Institutes of Health, Mammalian Gene Collection (MGC)
            unpublished (1999)
            Contact: Robert Strausberg, Ph.D.
            Email: cga@rs-remail.nih.gov
            Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
            cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiroki
            Toshiyuki and Piero Carninci (RIKEN)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LNLN at:
            http://image.llnl.gov
            Plate: LNA10678 row: 1 column: 23
            High quality sequence stop: 630.
            Location/Qualifiers
                1..630
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_image="479558"
                /clone_lib="NIH_MGC_96"
                /tissue_type="hypothalamus"
                /lab_host="DH10B"
                /note="Organ: brain; Vector: pBluescript (modified
                pBluescript KS+); Site: 1: BamHI; Site 2: SalI; XhoI (gtcag
                size-selected for average insert size 2.3 kb and
                normalized to 10^5. This is a primary library enriched
                for full-length clones and constructed using the
                Cap-trapper method (Carninci, in preparation) library
                constructed by M. Brownstein (NIH/NHGRI, National
                Institutes of Health). Note: this is a NIH_MGC library."

```

FEATURES

Source

1..630
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="479558"
/clone_lib="NIH_MGC_96"
/tissue_type="hypothalamus"
/lab_host="DH10B"
/note="Organ: brain; Vector: pBluescript (modified pBluescript KS+); Site: 1: BamHI; Site 2: SalI; XhoI (gtcag size-selected for average insert size 2.3 kb and normalized to 10^5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation) library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC library."

BASE COUNT 93 a 215 c 257 g 65 t

ORIGIN

Query Match 44.1%; Score 607.2; DB 12; Length 630;
Best Local Similarity 99.5%; Pred. No. 7.1e-74;
Matches 609; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

```

QY      7  TTCTTTGAGAGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      19  TCTCTTGGAGAGTGTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      67  GAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      79  GAGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      127  GAGAGGTCCGAGAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      139  GAGAGGTCCGAGAGTACGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      187  GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      199  GCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      247  GCTGCTCTCTGAGCGCTGGGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG
Db      259  GCTGCTCTCTGAGCGCTGGGCGCGCGCTGCGCGCGCGCGCGCGCGCGCG
QY      307  CTGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      319  CTGCGCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      367  GCGCGCGCGAGTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      379  GCGCGCGGAGTGGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      427  GCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      439  GCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      487  ACCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      499  ACCCGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      547  CGCGAAGCGGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
Db      559  CGCGAAGCGGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG
QY      607  CGTGTGTGTGAA 618
Db      619  CGTGTGTGTGAA 630

```

RESULT 12

BM893587/c

LOCUS

DEFINITION

BM893587 1j26g10.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

BM893587 1j26g10.x1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
cDNA clone IMAGE:6136123 3' similar to TR:Q9UEX4 Q9UEX4
HYPOPHYSEAL 22.8 KD PROTEIN ;, mRNA sequence.
BM893587
EST.
BM893587.1 GI:19349055
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 596)
Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K.,
Lemishka, I., Scearce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
Hillier, L., Marra, M., Page, D., Wyllie, T., Martin, J., Biststein, A.,
Schmitt, A., Theisling, B., Ritter, E., Ronko, I., Bennett, D., Cardenas,
M., Gibbons, M., McCann, R., Cole, R., Tsagaris, V., Williams, T.,
Jackson, Y., and Bowers, Y.
Endocrine Pancreas Consortium

JOURNAL
COMMENT

Unpublished (2000)
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: Juliana Brown
(brownjfas.harvard.edu)
Seq primer: -40up from Gbco
High quality sequence stop: 441.
Location/Qualifiers
1..596

FEATURES

99

BASE COUNT	139 a	183 c	169 g	105 t
ORIGIN				

Query Match	43.38;	Score 596;	DB 14;	Length 596;
Best Local Similarity	100.08;	Pred. No. 2.5e-72;		
Matches 596;	Conservative	0. Microbial		

QY	752	GCCTCTTTCTTCGAATTTTTTCGGGGGGGTGGCCCAAGCCAGCCCTCCCTCTCGGGGGGGCC	811
Db	596	GCCTTTTCTTCGAATTTTTTCGGGGGGGTGGCCCAAGCCAGCCCTCCCTCTCGGGGGGGCC	537
QY	812	ATGTGAGGCGTGAGACCTGAGACCAATGTGGGTGCGAGTGGTGTGGTGTACTACTT	871
Db	536	ATGTGAGGCGTGAGACCTGAGACCAATGTGGGTGCGAGTGGTGTGGTGTACTACTT	477
QY	872	GGCATCTATGCCAGCATCACAAGACAGCACCCTTCCGGATTTCGTGTACTCCGAC	931
Db	476	GGCATCTATGCCAGCATCACAAGACAGCACCCTTCCGGATTTCGTGTACTCCGAC	417
QY	932	TGGACAGACCTCCCGAGTCTTGCTTAGTGGCCCATGCAAACTGAGCTCATCTCTCACTC	991
Db	416	TGGACAGACCTCCCGAGTCTTGCTTAGTGGCCCATGCAAACTGAGCTCATCTCTCACTC	357
QY	992	CTAGACGAGGCGTGTGAGCGCTGACACCAAGCTCATCCAGAGAGCGCTGACCCCTCGAAT	1051
Db	356	CTAGACGAGGCGTGTGAGCGCTGACACCAAGCTCATCCAGAGAGCGCTGACCCCTCGAAT	297
QY	1052	ATTGTGAATGACTGTGGAGCGTGGGGGTAGAGCACTCTCCGTCTGCTCTGGCAAGAAATG	1111
Db	296	ATTGTGAATGACTGTGGAGCGTGGGGGTAGAGCACTCTCCGTCTGCTCTGGCAAGAAATG	237
QY	1112	GGAACACTGCGTCTCTCCGATCAGAGTGTGGCAGATGGGACAGCGGTGGATTTCTGCC	1171
Db	236	GGAACACTGCGTCTCTCCGATCAGAGTGTGGCAGATGGGACAGCGGTGGATTTCTGCC	177

[illegible]

RESULT 13

BM544255	LOCUS	596 bp	mRNA	linear	EST 20-FEB-2002
BM544255	DEFINITION	AGENCECOU7 6490655 NIH_MGC_125 Homo sapiens	CDNA	clone	IMAGE:5587759
BM544255	ACCESSION	5', URAA sequence.			
BM544255	VERSION	BM544255.1	GI:18775356		
BM544255	KEYWORDS	EST.			
BM544255	SOURCE	human.			
BM544255	ORGANISM	Homo sapiens			

FEATURES

10

Query Match 42.6%; Score 586.4; DB 13; Length 596;
Best Local Similarity 99.7%; Pred. No. 5e-71;
Matches 587; Conservative 0; Mismatches 2; Indels 0

789 CACCTTCGCTCTTCGGGGGGGGCCATGTTAGGCTGAGGCTTGAGCACAAGTGTGGTGC 848
 |||||
 1 CACCTTCGCTCTTCGGGGGGGGCCATGTTAGGCTGAGGCTTGAGCACAAGTGTGGTGC 848
 |||||
 849 AGGTGGGTGTGGGTGCTACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACTTCT 908
 |||||
 61 AGGTGGGTGTGGGTGCTACTACATTGGCATCTATGCCAGCATCAAGACAGACAGCACTTCT 120
 |||||
 909 CCGGATTTCGTGTGTAACGCCAGCTGGACAGCTCCCAAGTCTTTGAGTGCCCACTGC 968

DB 121 CCGGATTTCTGCTACTCCGACGCTCCAGCTCCAGCTCTTCTTACTGCCCACTGC 180
 QY 969 AAATGAGCTCATCTCTCTACTCTCTAGAGAGAGGTGTGAGGTGACACACAGCTATGC 1028
 DB 181 AAATGAGCTCATCTCTCTACTCTCTAGAGAGAGGTGTGAGGTGACACACAGCTATGC 240
 QY 1029 AGGAGGCTGCGCCCGCTGGAATATGTGTAATGATGAGAGGTGGGTGAGAGCACTCTC 1088
 DB 241 AGGAGGCTGCGCCCGCTGGAATATGTGTAATGATGAGAGGTGGGTGAGAGCACTCTC 300
 QY 1089 GCTCTGCTGCTGCGCAAGAAAGTGGAGAGTGGCTGTGCGATGAGTCTGGAGCAAG 1148
 DB 301 GCTCTGCTGCTGCGCAAGAAAGTGGAGAGTGGCTGTGCGATGAGTCTGGAGCAAG 360
 QY 1149 GGGCAGTGGCTGATTTCTGCGCCCAAGACAGAGAGTGTGTGTGCGCAAGTGAAGT 1208
 DB 361 GGGCAGTGGCTGATTTCTGCGCCCAAGACAGAGAGTGTGTGTGCGCAAGTGAAGT 420
 QY 1209 CCCCCAGTTGCTCTGCTCCAGAGAGCCAGAGTGGGTGCTCTCTGCTGCTCTCTCT 1268
 DB 421 CCCCCAGTTGCTCTGCTCCAGAGAGCCAGAGTGGGTGCTCTCTGCTGCTCTCTCT 480
 QY 1269 CTCTGATCTCTGCTCCAGAGAGCCAGAGTGGGTGCTCTCTGCTGCTCTCTCTCT 1328
 DB 481 CTCTGATCTCTGCTCCAGAGAGCCAGAGTGGGTGCTCTCTGCTGCTCTCTCTCT 540
 QY 1329 AATTAACCTAAGACCTCTATATATATATATATATATATATATATATATATATAT 1377
 DB 541 AATTAACCTAAGACCTCTATATATATATATATATATATATATATATATATATATAT 589

RESULT 14
 B1821899 941 bp mRNA linear EST 04-OCT-2001
 LOCUS 603035796F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5176967 5'
 DEFINITION B1821899 mRNA sequence.
 ACCESSION B1821899
 VERSION B1821899.1 GI:15933449
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 941)
 AUTHORS NIH-MGC http://mgi.ncl.nih.gov/
 TITLE Unpublished (1999)
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@bbs.fda.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: InCyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M11440 row: P column: 24
 High quality sequence stop: 779.
 Location/Qualifiers
 1. 941
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5176967"
 /clone_1="NIH_MGC_115"
 /lab="MGC-DH10B"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and

enriched for full-length clones and was constructed by C.
 Gruber (Invitrogen). Research Genetics tracking code
 021. Note: this is a NIH-MGC Library."
 BASE COUNT 130 a 320 c 329 g 162 t
 ORIGIN

Query Match 42.48; Score 583.2; DB 13; Length 941;
 Best Local Similarity 94.38; Pred. No. 1e-70;
 Matches 758; Conservative 0; Mismatches 28; Indels 18; Gaps 14;

QY 191 GAGGACCCCGGCGCTGCGGCTCCCGGAGGATGAGG-CCACCTCTGCTCTGCT 249
 DB 136 GAGGACCCCGGCGCTGCGGCTCCCGGAGGATGAGG-CCACCTCTGCTCTGCT 195
 QY 250 GCTCTGCTGCTGCGGCTGCGGCTCCCGGAGGATGAGG-CCACCTCTGCTCTGCT 309
 DB 196 GCTCTGCTGCTGCGGCTGCGGCTCCCGGAGGATGAGG-CCACCTCTGCTCTGCT 255
 QY 310 CCGGCGGACCCCGGCTGCGGCTCCCGGAGGATGAGG-CCACCTCTGCTCTGCT 368
 DB 256 CCGGCGGACCCCGGCTGCGGCTCCCGGAGGATGAGG-CCACCTCTGCTCTGCT 315
 QY 369 GCGCGATGCGCGGAGCGGCGGCGGAGCGGCGGCGG-CCACCTCTGCTCTGCT 427
 DB 316 GCGCGATGCGCGGAGCGGCGGCGGAGCGGCGGCGG-CCACCTCTGCTCTGCT 375
 QY 428 GCGGCGGAGCGGCGGAGCGGCGGCGGAGCGGCGGCGG-CCACCTCTGCTCTGCT 487
 DB 376 GCGGCGGAGCGGCGGAGCGGCGGCGGAGCGGCGGCGG-CCACCTCTGCTCTGCT 435
 QY 488 CCGGCGGAGCGGCGGAGCGGCGGCGGAGCGGCGGCGG-CCACCTCTGCTCTGCT 545
 DB 436 CCGGCGGAGCGGCGGAGCGGCGGCGGAGCGGCGGCGG-CCACCTCTGCTCTGCT 495
 QY 546 GCGGCGGAGCGGCGGAGCGGCGGCGGAGCGGCGGCGG-CCACCTCTGCTCTGCT 602
 DB 496 GCGGCGGAGCGGCGGAGCGGCGGCGGAGCGGCGGCGG-CCACCTCTGCTCTGCT 555
 QY 603 ACCGCGTGTGTGTAAGACAGAGGAGGACATTAAGAGCGGCGGCGGAGGTTACCTGCC 662
 DB 556 ACCGCGTGTGTGTAAGACAGAGGAGGACATTAAGAGCGGCGGCGGAGGTTACCTGCC 615
 QY 663 -AGTGTCTGCGGCTCTACTCTGCGGCTCA-TCGCAACCTCTACCGGCGGCGGAGGTT 720
 DB 616 GAGGTGCTGCGGCTCTACTCTGCGGCTCA-TCGCAACCTCTACCGGCGGCGGAGGTT 675
 QY 721 GTTGTATCTGTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
 DB 676 GTTGTATCTGTGTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 734
 QY 781 GCGGCGGAGCGGCGGCTGCGGCTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 839
 DB 735 GCGGCGGAGCGGCGGCTGCGGCTGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 794
 QY 840 TGTGGGTGCGGAGG-CCGCTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 898
 DB 795 TGTGGGTGCGGAGG-CCGCTGTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCT 853
 QY 899 AGCAGCTTCTCGGATTTCTGTGTACTCGAGCGGAGAGCGGCGGAGGAGGAGGAGG 958
 DB 854 AGCAGCTTCTCGGAG-TTCTGTGTACTCGAGCGGAGAGCGGCGGAGGAGGAGGAGG 912
 QY 959 TGCCCACTGCAAGAGTCAAGTCAAG 982
 DB 913 -CCCTGTCAAGTCAAGTCAAGTCAAG 933

RESULT 15
 B1821899 578 bp mRNA linear EST 29-APR-2002
 LOCUS 603035796F1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5136046 3'
 DEFINITION B1821899 mRNA sequence.
 ACCESSION B1821899
 VERSION B1821899.1 GI:15933449
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 578)
 AUTHORS NIH-MGC http://mgi.ncl.nih.gov/
 TITLE Unpublished (1999)
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: rgs@bbs.fda.gov
 Tissue Procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: InCyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L14M11440 row: P column: 24
 High quality sequence stop: 779.
 Location/Qualifiers
 1. 578
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:5136046"
 /clone_1="NIH_MGC_115"
 /lab="MGC-DH10B"
 /note="Organ: pooled brain, lung, testis; Vector:
 pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
 source anonymous pool of 6 male brains, age range 23-27; 1
 male lung, age 27; and 1 male testis, age 69. Library is
 oligo-dT primed and directionally cloned (EcoRV site is
 destroyed upon cloning). Average insert size 1.8 kb,
 insert size range 1-3 kb. Library is normalized and

Db	311	GGGCCACCATGACACCAAGGCGCTTGGCCGGGCGGATGTGGCCGACAGCGGCCGCGACGGCGC	370
OY	400	GCCCGGGGCTCCGGGAGAGAAAGCGAGCGCGGAGCGGGGACTTGGCGGGACTGTGAGG	459
Db	371	GGCCCGGGGCTCCGGGAGAGAAAGCGAGCGCGGGAGCGCGGACTGCCGGGACTCTGAGG	430
OY	460	GGACCCCGGGCGGGAGAGAGCGGGGACCCGCGGGGCCACCGCGGGCTCCGGGAGTG	519
Db	431	GGACCCCGGGCGGGAGAGAGCGGGGACCCGCGGGGCCACCGCGGGCTCCGGGAGTG	490
OY	520	CTCGGTGCTCTCCGGGATCCCGCTTACAGCGGCACAGCGCTCCGAGAGCGGGGTGCTCGGC	579
Db	491	CTCGGTGCTCTCCGGGATCCCGCTTACAGCGGCACAGCGCTCCGAGAGCGGGGTGCTCGGC	550
OY	580	GTCGAGAGCAACCTTTGGCCCTTCCAGCCCGCTGCTGTGTGACGAGCAGGACATTACAGCGC	639
Db	551	GTCGAGAGCAACCTTTGGCCCTTCCAGCCCGCTGCTGTGTGACGAGCAGGACATTACAGCGC	610
OY	640	GCTGACCGGCAAGTTACACTGCGCAGGTGCTCGGGGCTCTACTACTTGGCGGTCATGCCAC	699
Db	611	GCTGACCGGCAAGTTACACTGCGCAGGTGCTCGGGGCTCTACTACTTGGCGGTCATGCCAC	670
OY	700	CGCTACACCGGGCGAGCGCTGCAGTTTATCTGGTGAAGAAATGGCGAATCCATTGCCCTTT	759
Db	671	CGCTACACCGGGCGAGCGCTGCAGTTTATCTGGTGAAGAAATGGCGAATCCATTGCCCTTT	730
OY	760	CTTCCAGATTTTTCGGGGGTGGGCCCAAGCAGCCTGCTGTCGGGGGGGGGCGCATGTGAG	819
Db	731	CTTCCAGATTTTTCGGGGGTGGGCCCAAGCAGCCTGCTGTCGGGGGGGGGCGCATGTGAG	790
OY	820	GCTGGAGCGCTGAGAGACCAAGTGTGGGTGCAAGTGGGTGTGGGTGACACTATTGGACACTA	879
Db	791	GCTGGAGCGCTGAGAGACCAAGTGTGGGTGCAAGTGGGTGTGGGTGACACTATTGGACACTA	850
OY	880	TGCGACACATCAAGACAGACAGCACCCTTCCCGGATTTTGGGTGTACTCGACTGGGCACAG	939
Db	851	TGCGACACATCAAGACAGACAGCACCCTTCCCGGATTTTGGGTGTACTCGACTGGGCACAG	910
OY	940	CTGCCCAGCTTGTGGCTTAGTGGCCACTGCGAAGTGAAGTCAATGCTCTCACTCTAGAAGG	999
Db	911	CTGCCCAGCTTGTGGCTTAGTGGCCACTGCGAAGTGAAGTCAATGCTCTCACTCTAGAAGG	970
OY	1000	AGGCTGTAGGCTGTCAACAACCAAGTCATCAGAGAGGGGCGCCGCCCTGGATATTGGAA	1059
Db	971	AGGCTGTAGGCTGTCAACAACCAAGTCATCAGAGAGGGGCGCCGCCCTGGATATTGGAA	1030
OY	1060	TGACTAGAGGAGGTGGGGTATGAGAGCACTCCGCTCCCTGCTGCTGTGCGCAAGATGGACAGT	1119
Db	1031	TGACTAGAGGAGGTGGGGTATGAGAGCACTCCGCTCCCTGCTGCTGTGCGCAAGATGGACAGT	1090
OY	1120	GGCTGTCTGCGACATCAGTCTGGCAGCATATGGGGCAGTGGGTGATTTTGTGCCAAGACAG	1179
Db	1091	GGCTGTCTGCGACATCAGTCTGGCAGCATATGGGGCAGTGGGTGATTTTGTGCCAAGACAG	1150
OY	1180	AGGAGTGTGCTGTCTGGCAAGTAAAGTCCCGCAGTTTGTCTGTGTCACAGAGCCACAGG	1239
Db	1151	AGGAGTGTGCTGTCTGGCAAGTAAAGTCCCGCAGTTTGTCTGTGTCACAGAGCCACAGG	1210
OY	1240	TGGGCTGCTCTTTCCTGCTGCTCTGCTCTGCTGATCCTCCCGCACCCCTCTGCTGCTCT	1299
Db	1211	TGGGCTGCTCTTTCCTGCTGCTCTGCTCTGCTGATCCTCCCGCACCCCTCTGCTGCTCT	1270
OY	1300	GGGCGCGGCCCTTTTTCAGAGATCACTCAATAAAGCAAGCCTCATTAAGAAAAA	1359
Db	1271	GGGCGCGGCCCTTTTTCAGAGATCACTCAATAAAGCAAGCCTCATTAAGAAAAA	1330
OY	1360	AAAAAA 1366	
Db	1331	AAAAAA 1337	

```

? Sequence 1 Application US/09336536
? Patent No. 6406884
?
? GENERAL INFORMATION:
?
? APPLICANT: Leiby, K.
?
? APPLICANT: Bossone, S.
?
? TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF
?
? FILE REFERENCE: 7853-114
?
? CURRENT APPLICATION NUMBER: US/09/336,536
?
? CURRENT FILING DATE: 1999-06-18
?
? NUMBER OF SEQ ID NOS: 75
?
? SOFTWARE: PatentIn Ver. 2.0
?
? SEQ ID NO 1
?
? LENGTH: 1338
?
? TYPE: DNA
?
? ORGANISM: Homo sapiens
?
US-09-336-536-1

```

Query Match	94.28;	Score 1296.6;	DB 4;	Length 1338;
Best Local Similarity	99.68;	Pred. No. 7.7e-252;		
Matches 1310;	Conservative	0;	Mismatches 4;	Indels 1;

	Index	Gaps
54	CCAGGACTGGGGGAGACGGCCAGGGCCAGGGGGCCCTGGCCGGGGAGAAAGCCGGGGGCTGG	113
15	CCGGGACTGGGGGTGACGGAGGGCCAGGGGGCCCTGGCCGGGGAGAAAGCGCGGGGCTGG	74
114	AGCACCACCAACTGAGGGGTCCGGAGTACGAGACGGCCCCGAAGAGAGGGCCATCGGGGAGCC	173
75	AGCACCACCAACTGAGGGGTCCGGAGTACGAGACGGCCCCGAAGAGAGGGCCATCGGGGAGCC	134
174	GGGAGGGGGGACTGCGAGACGACCCCGGCTCCGGGCTCCCGGCTCCACAGCGCTATGAGGC	233
135	GGGAGGGGGGACTGCGAGAGGACCCCGGGCTCCGGGCTCCCGGCTCCAGCGCTATGAGGC	194
234	CACCTCTCTCTCTGCTGCTCTCTGCGGCTCGGCGCGCGCGCGCTCCCGCCCACTGAGCACA	293
195	CACCTCTCTCTCTGCTGCTCTCTGCGGCTCTGCGGCGCGCGCGCTCCCGCCCACTGAGCACA	254
294	AGATCCCGACCTCTGCGCGGGGACCCCGGCTTCCAGAGCAGCGCGGGCCACCATGACA	353
255	AGATCCCGACCTCTGCGCGGGGACCCCGGCTTCCAGAGCAGCGCGGGCCACCATGACA	314
354	GCCAGGGCTTGGCGGGCCGCGCATGGCCCGCGAGCGGGCGGAGCGGGCGCGCGGGCTCCGG	413
315	GCCAGGGCTTGGCGGGCCGCGCATGGCCCGCGAGCGGGCGGAGCGGGCGCGCGGGCTCCGG	374
414	GAGAAAAAGCGAGGGCGGGAGCGCGGGACTGCGCGGACCTCGAGGGGACCCCGGGCGCG	473
375	GAGAAAAAGCGAGGGCGGGAGCGCGGGAGG - GGGAGTGGCGGGACCTCGAGGGGACCCCGGGCGCG	433
474	GAGAGAGAGCGGGAGCCCGCGGGGGCCACCGGGCTCCCGGGGAGTGTCTGGTGGCTCCGC	533
434	GAGAGAGAGCGGGAGCCCGCGGGGGCCACCGGGCTCCCGGGGAGTGTCTGGTGGCTCCGC	493
534	GATCGGCTTCAAGCGCAAGCGCTCCGAGAGCGCGGTCTCCGCGCGCTGAGAGCAACCT	593
494	GATCGGCTTCAAGCGCGCAAGCGCTCCGAGAGCGCGGTCTCCGCGCGCTGAGAGCAACCT	553
594	TGCGCTTCAAGCGCGGTCTGTGAGCAGCAGGACATTACAGCGCGCTACCGGCAAGT	653
554	TGCGCTTCAAGCGCGGTCTGTGAGCAGCAGGACATTACAGCGCGCTACCGGCAAGT	613
654	TCACCTCCAGAGTCCGCGGCTGTACTACTTGGCGGTCATGCAACCGCTTACCGGGCCA	713
614	TCACCTCCAGAGTCCGCGGCTGTACTACTTGGCGGTCATGCAACCGCTTACCGGGCCA	673
714	GCCGCGAGTTGATCTGGTGAAGAAATGCGGAATCCATTGCTCTTCTTCCAGTTTTCG	773
674	GCCGCGAGTTGATCTGGTGAAGAAATGCGGAATCCATTGCTCTTCTTCCAGTTTTCG	733
774	GGGGGTGGCCCAAGCCAGCGCTGCTCGGGGGGGGCCCATGTGAGAGCTGAGGCTTGAG	833
734	GGGGGTGGCCCAAGCCAGCGCTGCTCGGGGGGGGCCCATGTGAGAGCTGAGGCTTGAG	793


```

RESULT 5
US-09-188-930-218
: Sequence 218, Application US/09188930A
: Patent No. 6150502
:
: GENERAL INFORMATION:
:
: APPLICANT: Watson, James D.
: APPLICANT: Strachan, Lorna
: APPLICANT: Sleeman, Matthew
: APPLICANT: Onrust, Rene
:
: TITLE OF INVENTION: Compositions Isolated From Skin Cells
: TITLE OF INVENTION: and Methods For Their Use
: FILE REFERENCE: 11000.1011c1
: CURRENT APPLICATION NUMBER: US/09/188,930A
: CURRENT FILING DATE: 1998-11-09
: NUMBER OF SEQ ID NOS: 348
:
: SOFTWARE: FastSeq for Windows Version 3.0
:
: SEQ ID NO 218
:
: LENGTH: 1001
:
: TYPE: DNA
:
: ORGANISM: Rat
:
: US-09-188-930-218

```

Query Match	44.0%	Score 606;	DB 3;	Length 1001;
Best Local Similarity	78.6%;	Pred. No. 2.7e-113;		
Matches 739; Conservative	0;	Mismatches 195;	Indels 6;	Gaps 1;

QY 66 TGACGGCAGCGCCAGGGGGCCCTCGCCCGGGGAGAAACCGGGGGCTGGAGCACCAAC 125
 Db 31 TGTACAGAGAGGGAGGGCTGCTGGTCTGGTGGGTAGAGAGGAGGAGCGCCAGCGAGG 90
 QY 126 TGGAGGGTCCGAGTAGCGAGCGCCCGAAGAGAGGCATTCGGGGAGCCGGCAGGGGGAC 185
 Db 91 GTCTGAGGAAGGCATTCAGAGGAGGAGGAGCTGGGAGAGCTGGGAGACCGGGAAGGGCTTAC 150
 QY 186 TG-----CGACAGCAACCCCGGCGTCCGGGGCTCCGGTGGCCAGGCGTATGAGGCACACTCC 239
 Db 151 AGCTACAGAAAGAGATCTGGCGCTCTGGGGCTCTCTGGTCAATCATTGAGGCCACTTC 210
 QY 240 TCGTCTCGTGTCTCTGGGGCCTCGGGGGCGGGTGGCCCCCTGAGAGACAACAAATCC 299
 Db 211 TTGCCCTGTGCTTCTGGGTCTGGCATAGGCTCTCTCTGTGGAGACAAAGAATCC 270
 QY 300 CCAGCTCTGCCCGGGGACCCCGGCTTCCAGGACGCGCGGCGCACATGAGCAGCAGG 359
 Db 271 CCAGGCTGTGCCCGGGACGCCCGCTCCAGGACACAGGCGCACACAGCGCAGCCAAAG 330
 QY 360 GCTTCCCGGCGCGCATGTGGCGGACGCGCCGACGCGCGCCCGGGGCTTCCGGAGAGA 419
 Db 331 GCTCTCCCTGGCGCTGACGCGCGTATGGCGCGACGGTGGACCGCGAGTCTCCGGGAGAGA 390
 QY 420 AAGCGAGAGGCGGAGAGCGCGGACTGGCGGACTTCAGAGGAGACCCGGGCGCGAGAGG 479
 Db 391 AAGCGAGAGGCGGAGAGCCGGACTTACTCTGGGCCAGCTGGGAGACCCGGGCCCGCTGGAG 450
 QY 480 AGCGGGGACCCGGGGGGCCACCGGGGCTCCCGGGGAGTGTCTCGTGGCTCCGCGATCCG 539
 Db 451 AGCGAGGACCTGTGGGGCTATGGCGGCTCCGGGGGAGTGTCTGGTGGCCCCACGATCAG 510
 QY 540 CTTTACAGCGCAAGCGCTCCGAGAGCGCGGTGGCTCGCGCGTGTGAGAGCAACCTTCCT 599
 Db 511 CTTTCACTGCCAAGCGCATCAGAGAGCGGGTACCTCTCCGACCGGACACACCCCTTACCT 570
 QY 600 TCGACCGCGTGGTGGTGAACGAGCAGGAGCATTTAGAGCCCGCTACACGGCAAGTTACCT 659
 Db 571 TCGACCGTGGTGGCTGAAGAGCAGAGGACATTACGATCCCACTACCGGCAAGTTACCT 630
 QY 660 GCCAGGTGCTGGGGTCTACTACTTGGCGCTCATCGACCGGTAAACCGGGCAGCCTGC 719
 Db 631 GCCAAGTGCCTGGTGTACTACTTGGCTGTCAATCCCACTGTCTTACCGGGCAGCCTAC 690
 QY 720 AGTTGATGTGGTGAAGATGGCGAATTCATTCGCTTTCTTCCAGTTTTCGGGGGGT 779

QY	227	ATGAGCCACACCTCCTGCTCCTGCCTCCTCGGCGCTTGCGGGGCTGCCCGGCTGCCCGCCACTGAGC	286
Db	1	ATGMCNCCNYNTNGTNYTYNTYTNTGNGANTNCNCNGMNGMSNCCNCTGTGAY	60
QY	287	GACAACAGATCCCAGCCTTCGTGCCCCGGGACCOCGGCCCTCCAGCACACCCGGGCCAC	346
Db	61	GAYAAYAAARATHCCMNSYNTTGTCCNGMGATTCGNGANTTCCNGSNMACNCCNGNCAY	120
QY	347	CATGGCAGCCAGGGCTTGC CGGGCCGCGATGCGCGCAGACCGCCGCAACGCGCGCCCGG	406
Db	121	CAYGMNSNARGNYNTNCCNGNMNGAYGVMMNGAYGNMNGAATYGNNGNATYGNNGCNCCNGN	180
QY	407	GCTCCGGAGAGAAGAGCGGGCGGGGAGCGCCGGAGCTCCGGGACCTGAGGGGACCCC	466
Db	181	GCNCCNGNBARAARGNNGARGNGMGKCCNGGYTTCCNCGNCCNMGNGNGAYCCN	240
QY	467	GGGCGCGGAGAGAGCGGGGACCCCGCGGGCCACCAGCGGCTGCCGGGAGTGTCCGTG	526
Db	241	GGNCCNMNNGNBARCNGNCCNGCNGCNAAGNCCNCGNCGNARGNARTGYSMNGTN	300
QY	527	CCTCGCGGATCGCGCTTGAAGCGCCAAGCCGCTCCGAGACCGCGGGTCCCTCGCGCTGAC	586
Db	301	CCNCCMKNMNSNCGNTTWSNCGNAARMGMSNGARMSNMNGTNCNCCNCCNMSNGAY	360
QY	587	GCAACCTTGCCCTTCGACCGCGTGTGTAAGACAGACAGCATTAGACGCGCTCAC	646
Db	361	GCNCCNYTNCNTTGYAYNGNTNTNGTNAIYGARCARGNCAITTAGATGCGNTNACN	420
QY	647	GGCAATTACCTCGCCAGGTGCGTGGGCTCTACTACTTTCGCGCATGCCACCTCTAC	706
Db	421	GGMAATTACTGTGCARTCCNCGNNTMTAVTYTYYGCGTNCAGACNACNTNTAY	480
QY	707	CGGGCCAGGCTCAGTTGATGTGTGAAGATGSCGATCATGCGCTTCTCTCCAG	766
Db	481	MONGNMSNYTNCARTTGYAYTNGTNAARAAYGNGARGMSATHGCMNSTTTTTYCAR	540
QY	767	TTTTTCGGGGGGTGGCCCCAACGCACTCGCTCGGGGGGCCATGTAGAGCTGAG	826
Db	541	TTTTTGGNGTGGCCNAARCCNCSNYSNTNMSGNCGNCGATGTWNGVTNCR	600
QY	827	CCTGAGGACCAAGTGTGCTCAGGTGGGTGGGTGACTAATTGGCATCTATGCCAGC	886
Db	601	CCNGARGAYCARGTWTGGTNCANGSTNGNTNGNAGYTYATHGGVATHHTAYGCMWSN	660
QY	887	ATCAAAGACAGACAGCACTTCTCCGGATTTCTGTGTACTCCGACTGCGACAGCTCCCA	946
Db	661	ATHAARACAGYWSNACNTTYSNGNTTYTNGTNTAYMSNGAYTGCAYWMSMWSN	720
QY	947	GCTCTTGC 954	
Db	721	GTNTTTCG 728	
RESULT 9			
US-09-336-536-74			
Sequence 74, Application US/09336536			
Patent No. 6406884			
GENERAL INFORMATION:			
APPLICANT: Leidy, K.			
APPLICANT: McKay, C.			
APPLICANT: Bossone, S.			
TITLE OF INVENTION: SECRETED PROTEINS AND USES THEREOF			
FILE REFERENCE: 7853-144			
CURRENT APPLICATION NUMBER: US/09/336,536			
CURRENT FILING DATE: 1999-06-18			
NUMBER OF SEQ. ID NOS: 75			
SOFTWARE: PatentIn Ver. 2.0			
SEQ. ID NO. 74			
LENGTH: 601			
TYPE: DNA			
ORGANISM: Rattus norvegicus			
FEATURE:			

NAME/KEY: modified_base
LOCATION: all "n" positions
OTHER INFORMATION: n-a, c, g, or t
US-09-336-536-74

Query Match 18.7% Score 257.8; DB 4; Length 601;
Best Local Similarity 72.0%; Pred. No. 1.8e-43;
Matches 365; Conservative 0; Mismatches 134; Indels 8; Gaps 2;

QY 66 TGACGGCAGGCGGCGGCGCTGGCCGGGAGAAAGCCGGGGCTGGAGCACCACCAAC 125
DB 96 TGTACGAGGCGGCGGCGCTGGCTGGGTTAGAGATGGAGACAGGCGCCAGAGAGG 155
QY 126 TGGAGGCTCCGAGTAGAGAGAGCCCGGAGAGAGCCATCGGGAGCCGGAGGGGAG 185
DB 156 GTCTAGAGAGGCAATTCAAGCAGACAGCTGGAGAGCTGGGAGAGCCGGAAAGGCTTAC 215
QY 186 TG-----CGAGAGACCCCGGCGCTCCGGCTCCGAGCTATGAGGCACTTC 239
DB 216 AGACTACAAAGAGAGATCTGGCGCTCTGGGCTATCAGCATAGAGCCACTTC 275
QY 240 TGTCTCTGCTGCTCTGCGGCTGGCGGCGGCTGCCCCACTGAGACAGCAAGATCC 299
DB 276 TTGGCCCTGCTCTCTGAGGCTGGCATAGAGCTCTCTCTCTGAGACAGCAAGATCC 335
QY 300 CCAGCCTCTGCGCGGCGGCGGCTTCCAGGACAGCGGGGCAACATGGAGAGCCAG 359
DB 336 CCAGCCTGTGCTCCGCGGCGGCGGCTCCAGGACAGCCAGCAACAGGCGGCAAG 395
QY 360 GCTTGGCGGCGGCGGATGCGCGGCGGCGGCGGCGGCGGCTCCGGAGAGA 419
DB 396 GCTTGGCGGCGGCGGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGAGA 455
QY 420 AAGCGGAGCGGCGGAGCGGCGGAGCTGCGGAGCTGAGAGGAGCCCGGGCGCGAGAG 479
DB 456 AAGCGGAGCGGCGGAGCGGCGGAGCTGCGGAGCTGAGAGGAGCCCGGGCGCGAGAG 515
QY 480 AGCGGAGCGGCGGCGGCGGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTCG 539
DB 516 AGCGGAGCGGCGGCGGCGGAGCTGCGGAGCTGCGGAGCTGCGGAGCTCGGAGCTCG 573
QY 540 CCTTACGCGGCGGCGGCGGAGAGCC 566
DB 574 GCTTACGCGGCGGCGGAGAGAGCC 600

RESULT 10
US-09-188-930-26

Sequence 26, Application US/09188930A
Patent No. 6150502
GENERAL INFORMATION:
APPLICANT: Watson, James D.
APPLICANT: Strachan, Lorna
APPLICANT: Sleeman, Matthew
APPLICANT: Murlison, James Greg
TITLE OF INVENTION: Compositions Isolated From Skin Cells
FILE REFERENCE: 11000.1011C1
CURRENT APPLICATION NUMBER: US/09/188.930A
NUMBER OF SEQ ID NOS: 348
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 26
LENGTH: 393
TYPE: DNA
ORGANISM: Rat
US-09-188-930-26

Query Match 10.9% Score 149.6; DB 3; Length 393;
Best Local Similarity 69.1%; Pred. No. 8.6e-22;
Matches 250; Conservative 0; Mismatches 104; Indels 8; Gaps 3;

QY 66 TGACGGCAGGCGGCGGCGCTGGCCGGGAGAAAGCCGGGGCTGGAGCACCACCAAC 125
DB 33 TGTACGAGGCGGCGGCGCTGGCTGGGTTAGAGATGGAGACAGGCGCCAGAGAGG 92
QY 126 TGGAGGCTCCGAGTAGAGAGAGCCCGGAGAGAGCCATCGGGAGCCGGAGGGGAG 185
DB 93 GTCTAGAGAGGCAATTCAAGCAGACAGCTGGAGAGCTGGGAGAGCCGGAAAGGCTTAC 152
QY 186 TG-----CGAGAGACCCCGGCGCTCCGGCTCCGAGCTATGAGGCACTTC 239
DB 153 AGACTACAAAGAGAGATCTGGCGCTCTGGGCTATCAGCATAGAGCCACTTC 272
QY 240 TGTCTCTGCTGCTCTGCGGCTGGCGGCGGCTGCCCCACTGAGACAGCAAGATCC 299
DB 213 TTGGCCCTGCTCTCTGAGGCTGGCATAGAGCTCTCTCTCTGAGACAGCAAGATCC 335
QY 300 CCAGCCTGTGCGCGGCGGCGGCTTCCAGGACAGCGGGGCAACATGGAGAGCCAG 359
DB 273 CCAGCCTGTGCTCCGCGGCGGCGGCTCCAGGACAGCCAGGCGGCGGCAAG 392
QY 360 GCTTGGCGGCGGCGGATGCGCGGCGGCGGCGGCGGCGGCGGCTCCGGAGAGA 419
DB 333 GCTTGGCGGCGGCGGATGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGAGA 455
QY 420 AA 421
DB 391 AA 392

RESULT 11
US-08-463-911-1

Sequence 1, Application US/08463911
Patent No. 5869330
GENERAL INFORMATION:
APPLICANT: Scherer, Philipp E.
APPLICANT: Lodish, Harvey F.
TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Millitia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: USA
ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/463.911
FILING DATE:
CLASSIFICATION: 530
ATTORNEY/AGENT INFORMATION:
NAME: Granahan, Patricia
REGISTRATION NUMBER: 32,227
REFERENCE/DOCKET NUMBER: WH195-05
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 861-6240
TELEFAX: (617) 861-9540
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1276 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 46..786
US-08-463-911-1

Query Match 6.6%; Score 90.6; DB 2; Length 1276;
 Best Local Similarity 51.5%; Pred. No. 7.3e-10;
 Matches 316; Conservative 0; Mismatches 279; Indels 18; Gaps 4;

QY 333 GCACCGCCGCGCCACCATGACAGCCGCTTGCCTGCGCGCGGATGCGCGCAGCGCCGCG 392
 DB 170 GATGAGCAGGACATCCAGACATCTCTGACACAAATGCGCACACAGCGCGTATGAGAG 229
 QY 393 ACGCGCGCCGCGCGCGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 452
 DB 230 ATGGACACTCTGAG 289
 QY 453 CTCGAGGAG 512
 DB 290 AGACAGGAG 349
 QY 513 GCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 568
 DB 350 GCAG 409
 QY 569 GTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 620
 DB 410 TCGAG 469
 QY 621 AGCAGGAG 680
 DB 470 AACGAG 529
 QY 681 ACTTGGCG 740
 DB 530 ACTTCTTACAG 589
 QY 741 GCGAG 800
 DB 590 ACAAG 846
 QY 801 CCGGCG 857
 DB 647 CTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 706
 QY 858 TGGGTACTGAG 917
 DB 707 ATGGGAG 766
 QY 918 TGGGTACTGAG 766
 DB 767 TTCTTACATGGA 779

RESULT 12

US-08-463-911-6
 Sequence 6, Application US/08463911
 Patent No. 5869330
 GENERAL INFORMATION:
 APPLICANT: Scherer, Philipp E.
 APPLICANT: Lodish, Harvey F.
 TITLE OF INVENTION: A NOVEL SERUM PROTEIN PRODUCED
 TITLE OF INVENTION: EXCLUSIVELY IN ADIPOCYTES
 NUMBER OF SEQUENCES: 7
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
 STREET: Two Millia Drive
 CITY: Lexington
 STATE: Massachusetts
 COUNTRY: USA
 ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463,911
 FILING DATE:
 CLASSIFICATION: 530
 ATTORNEY/AGENT INFORMATION:
 NAME: Granahan, Patricia
 REGISTRATION NUMBER: 32,227
 REFERENCE/DOCKET NUMBER: WH195-05
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (617) 861-6240
 TELEFAX: (617) 861-9540
 INFORMATION FOR SEQ ID NO: 6:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1313 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 73..804
 US-08-463-911-6

Query Match 6.0%; Score 82.2; DB 2; Length 1313;
 Best Local Similarity 50.1%; Pred. No. 3.6e-08;
 Matches 316; Conservative 0; Mismatches 303; Indels 12; Gaps 4;

QY 321 CCGGCTTCCAG 380
 DB 194 CCGGCTTCCAG 253
 QY 381 CCGGCG 440
 DB 254 CCGGCTGAG 313
 QY 441 GACTCCCGGAG 500
 DB 314 GTGAAG 373
 QY 501 CCGGCG 560
 DB 374 AAGGAG 430
 QY 561 AAGGCG 620
 DB 431 AGACTTACGTTACTATCC--CAACATGCCGATTCGCTTACAGATCTTCTACATC 487
 QY 621 AGCAGGAG 680
 DB 488 AGCAAAACACTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 547
 QY 681 ACTTGGCG 740
 DB 548 ACTTGGCTTACAGATCACTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 607
 QY 741 CCGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 800
 DB 608 ACAAGGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 859
 QY 801 CCGGCG 859
 DB 665 CCGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 724
 QY 860 --GGTACTACATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 917
 DB 725 AAGGAG 784
 QY 918 TGGGTACTGAG 948
 DB 785 TTCTTACATGAG 815

RESULT 13
 US-09-140-804-9

```

; Sequence 9, Application US/09140804
; Patent No. 6197930
; GENERAL INFORMATION:
; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-49
; CURRENT APPLICATION NUMBER: US/09/140,804
; EARLIER FILING DATE: 1998-08-26
; EARLIER APPLICATION NUMBER: 60/056,983
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 9
; LENGTH: 4517
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-140-804-9

```

```

Query Match      6.0%; Score 82.2; DB 4; Length 4517;
Best Local Similarity 50.1%; Pred. No. 4.4e-08;
Matches 316; Conservative 0; Mismatches 303; Indels 12; Gaps 4;

```

```

QY 321 CCGGCTTCAGGACCGCCGCGCCAGCATGCGACGAGCGCTTCCCGGCGCGATGGCC 380
DB 148 CCGGATCCCGGCGCGATCCGCGCATATGCGGCGCCCGAGCGCGATGCGAGATGCA 207
QY 381 CGGAGCGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 440
DB 208 CCGGCTTCAGGACCGCCGCGCCAGCATGCGACGAGCGCTTCCCGGCGCGATGGCC 267
QY 441 GACTGCGCGGACCTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 500
DB 268 GTGAACCGGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 327
QY 501 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 560
DB 328 AAGGAGCACTGGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 384
QY 561 AGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 620
DB 385 AGACTTACGTTACTATCC---CCAGATGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 441
QY 621 AGGAGGACATTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 680
DB 442 AGGAGGACATTCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 501
QY 681 ACTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 740
DB 502 ACTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 561
QY 741 GCGAATTCATTCGCTTCTTCAGTCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 800
DB 562 ACAGAGCTATTCGCTTCTTCAGTCTTTCGCGCGCGCGCGCGCGCGCGCGCGCGCG 618
QY 801 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 859
DB 619 CCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 678
QY 860 --GATGACATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 917
DB 679 AAGGAGGAGTATTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 917
QY 918 TCGTGTACTCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 948
DB 739 TCTCTACCATGACACCAACTGATCACCAGCT 769

```

```

RESULT 14
US-09-118-408-23
; Sequence 23, Application US/09118408A
; Patent No. 6265544
; GENERAL INFORMATION:

```

```

; APPLICANT: Sheppard, Paul O.
; TITLE OF INVENTION: ADIPOCYTE-SPECIFIC PROTEIN HOMOLOGS
; FILE REFERENCE: 97-30
; CURRENT APPLICATION NUMBER: US/09/118,408A
; EARLIER FILING DATE: 1998-07-17
; EARLIER APPLICATION NUMBER: 60/053,154
; EARLIER FILING DATE: 1997-07-18
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FASTSEQ for Windows Version 3.0
; SEQ ID NO 23
; LENGTH: 843
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Degenerate nucleotide sequence encoding zs1937
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(843)
; OTHER INFORMATION: Each N is independently any nucleotide.
US-09-118-408-23

```

```

Query Match      5.6%; Score 77.4; DB 4; Length 843;
Best Local Similarity 30.0%; Pred. No. 3e-07;
Matches 209; Conservative 99; Mismatches 384; Indels 5; Gaps 2;

```

```

QY 233 CCACTCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTGCTCTG 292
DB 131 CCMNSCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCNCCN 190
QY 293 AAGATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 352
DB 191 AAGATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 250
QY 353 AGCCAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 410
DB 251 AATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 310
QY 411 CCGGAGAGAAAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 470
DB 311 GNGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 370
QY 471 CCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 530
DB 371 AATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 430
QY 531 CCGGATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 590
DB 431 AATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 490
QY 591 CCGGATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 650
DB 491 CCGGATCCCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 550
QY 651 AGTTACCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 707
DB 551 ARTTYATATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 610
QY 708 GGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 767
DB 611 AATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 670
QY 768 TTTTCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 827
DB 671 ARGNGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 730
QY 828 CTGAGAGCAAGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 887
DB 731 AATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 790
QY 888 TCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 924
DB 791 TNGAYACNTATATACNTATATACNTATATACNTATATACNTATATACNTATATAC 827

```


GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 22, 2003, 04:43:36 ; Search time 225 Seconds
(without alignments)
8980.654 Million cell updates/sec

Title: US-09-943-851A-41

Perfect score: 1377
1 gactgtctctcttgagctc.....aaaaaaaaaaaaaaaa 1377

Scoring table:
IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1042519 seqs, 733713590 residues

Total number of hits satisfying chosen parameters: 2085038

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

Published Applications, NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/ECT_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/ECTUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1377	100.0	1377	9	US-09-944-413-41
2	1377	100.0	1377	9	US-09-944-403-41
3	1377	100.0	1377	9	US-09-944-896-41
4	1377	100.0	1377	9	US-09-944-944-41
5	1377	100.0	1377	9	US-09-944-907-41
6	1377	100.0	1377	9	US-09-944-929-41
7	1377	100.0	1377	9	US-10-028-072-361
8	1377	100.0	1377	9	US-10-121-048-361
9	1377	100.0	1377	9	US-10-123-904-361
10	1377	100.0	1377	9	US-10-140-470-361
11	1377	100.0	1377	9	US-10-175-746-361
12	1377	100.0	1377	9	US-10-176-918-361
13	1377	100.0	1377	9	US-10-176-921-361
14	1377	100.0	1377	9	US-10-137-865-361
15	1377	100.0	1377	9	US-10-140-474-361
16	1377	100.0	1377	9	US-10-142-431-361
17	1377	100.0	1377	9	US-10-143-114-361
18	1377	100.0	1377	9	US-10-140-002-361
19	1377	100.0	1377	9	US-10-142-419-361

20	1377	100.0	1377	9	US-10-123-262-361	Sequence 361, App
21	1377	100.0	1377	9	US-10-142-423-361	Sequence 361, App
22	1377	100.0	1377	9	US-10-121-050-361	Sequence 361, App
23	1377	100.0	1377	9	US-10-141-755-361	Sequence 361, App
24	1377	100.0	1377	9	US-10-143-032-361	Sequence 361, App
25	1377	100.0	1377	9	US-10-123-108-361	Sequence 361, App
26	1377	100.0	1377	9	US-10-123-236-361	Sequence 361, App
27	1377	100.0	1377	9	US-10-123-261-361	Sequence 361, App
28	1377	100.0	1377	9	US-10-140-921-361	Sequence 361, App
29	1377	100.0	1377	9	US-10-140-928-361	Sequence 361, App
30	1377	100.0	1377	9	US-10-121-045-361	Sequence 361, App
31	1377	100.0	1377	9	US-10-123-292-361	Sequence 361, App
32	1377	100.0	1377	9	US-10-123-903-361	Sequence 361, App
33	1377	100.0	1377	9	US-10-124-819-361	Sequence 361, App
34	1377	100.0	1377	9	US-10-124-822-361	Sequence 361, App
35	1377	100.0	1377	9	US-10-140-925-361	Sequence 361, App
36	1377	100.0	1377	9	US-10-160-498-361	Sequence 361, App
37	1377	100.0	1377	9	US-09-944-884-41	Sequence 41, App1
38	1377	100.0	1377	9	US-10-121-041-361	Sequence 361, App
39	1377	100.0	1377	9	US-10-121-043-361	Sequence 361, App
40	1377	100.0	1377	9	US-10-121-047-361	Sequence 361, App
41	1377	100.0	1377	9	US-10-123-215-361	Sequence 361, App
42	1377	100.0	1377	9	US-10-123-902-361	Sequence 361, App
43	1377	100.0	1377	9	US-10-123-908-361	Sequence 361, App
44	1377	100.0	1377	9	US-10-123-909-361	Sequence 361, App
45	1377	100.0	1377	9	US-10-123-910-361	Sequence 361, App

ALIGNMENTS

RESULT 1
US-09-944-413-41
Sequence 41, Application US/09944413
Patent No. US20020156004A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gertsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavio, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P2548P1
CURRENT APPLICATION NUMBER: US/09/944,413
PRIOR FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067,411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,696
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997

QY 1141 GCAGATGGGGCAGTGGCTGATTTCTGCCCAACAGCAGAGAGTGTCTGTGCGCAA 1200
Db 1141 GCACATAGGGGGCAGTGGCTGATTTCTGCCCAACAGCAGAGAGTGTCTGTGCGCAA 1200
QY 1201 GTGTAAGTCCCGCAGTGTCTGTGTCAGAGAGCCAGCGTGGGTGCTCTTCTGTGTC 1260
Db 1201 GTGTAAGTCCCGCAGTGTCTGTGTCAGAGAGCCAGCGTGGGTGCTCTTCTGTGTC 1260
QY 1261 CTCTGCTTCTGTGATCTCTCCCAACCCCTCTCTGCTGTGGGGCCGCTTTTCTGAGA 1320
Db 1261 CTCTGCTTCTGTGATCTCTCCCAACCCCTCTCTGCTGTGGGGCCGCTTTTCTGAGA 1320
QY 1321 GATCAGTCAATTAACCTAAGAACCCCTATATATATATATATATATATATATATATAT 1377
Db 1321 GATCAGTCAATTAACCTAAGAACCCCTATATATATATATATATATATATATATATAT 1377

RESULT 2
US-09-944-403-41
Sequence 41. Application US/09944403
Patent No. US20020165143A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Eaton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Filvaroff, Ellen
APPLICANT: Gerltsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaldi, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillan, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumas, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2548P1C1
CURRENT APPLICATION NUMBER: US/09/944,403
CURRENT FILING DATE: 2001-09-26
PRIOR APPLICATION NUMBER: 09/866,028
PRIOR FILING DATE: 2001-05-25
PRIOR APPLICATION NUMBER: 60/067,411
PRIOR FILING DATE: December 3, 1997
PRIOR APPLICATION NUMBER: 60/069,334
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,335
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,278
PRIOR FILING DATE: December 11, 1997
PRIOR APPLICATION NUMBER: 60/069,425
PRIOR FILING DATE: December 12, 1997
PRIOR APPLICATION NUMBER: 60/069,966
PRIOR FILING DATE: December 15, 1997
PRIOR APPLICATION NUMBER: 60/069,694
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,702
PRIOR FILING DATE: December 16, 1997
PRIOR APPLICATION NUMBER: 60/069,870
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/069,873
PRIOR FILING DATE: December 17, 1997
PRIOR APPLICATION NUMBER: 60/068,017
PRIOR FILING DATE: December 18, 1997
PRIOR APPLICATION NUMBER: 60/070,440
PRIOR FILING DATE: January 5, 1998
PRIOR APPLICATION NUMBER: 60/074,086
PRIOR FILING DATE: February 9, 1998
PRIOR APPLICATION NUMBER: 60/074,092

QY 1 GACTAGTCTCTTGAGTCTGTGAGAGAGAAAGCGAGCCGCGAGGAGCGAAGCAGAGAC 60
Db 1 GACTAGTCTCTTGAGTCTGTGAGAGAGAAAGCGAGCCGCGAGGAGCGAAGCAGAGAC 60
QY 61 TGGGCTGACGCGAGGCGAGGCGGCGCTGCGCGGGGAGAGACCGCGGGGCTGAGACACA 120
Db 61 TGGGCTGACGCGAGGCGAGGCGGCGCTGCGCGGGGAGAGACCGCGGGGCTGAGACACA 120
QY 121 CCAACTGAGGGTCCGAGTAGCGAGCCGCCCGGAAAGAGGCCATCGGGAGACCGGAGAGG 180
Db 121 CCAACTGAGGGTCCGAGTAGCGAGCCGCCCGGAAAGAGGCCATCGGGAGACCGGAGAGG 180
QY 181 GGGAGTGGAGAGAGACCCCGCGGCTCCGCGGTCCAGCCCTATGAGGCACTCTCT 240
Db 181 GGGAGTGGAGAGAGACCCCGCGGCTCCGCGGTCCAGCCCTATGAGGCACTCTCT 240

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
US-09-944-403-41
TYPE: DNA
ORGANISM: Homo Sapien
SEQ ID NO 41
LENGTH: 1377
PRIOR APPLICATION NUMBER: PCT/US00/08439
PRIOR FILING DATE: March 30, 2000
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: May 22, 2000
PRIOR APPLICATION NUMBER: PCT/US00/20710
PRIOR FILING DATE: July 28, 2000
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: December 1, 2000
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: February 28, 2001
NUMBER OF SEQ ID NOS: 120

QY	241	CGTCTCTGTCTCTCTGGGCGCTTGGCGCGCGGCTCGCCCCCTACATGGAGACAACAAGATCCC	300
Db	241	CGTCTCTGTCTCTCTGGGCGCTTGGCGCGCGGCTCGCCCCCTACATGGAGACAACAAGATCCC	300
QY	301	CAGCCTCTGCCCGGGGACACCCCGGCTTTCAGAGCAGCGCGGCACCATATGGACACAGG	360
Db	301	CAGCCTCTGCCCGGGGACACCCCGGCTTTCAGAGCAGCGCGGCACCATATGGACACAGG	360
QY	361	CTTCCCGGGCGCGCATGATGGCGCGACAGGCGCGCGAGCGCGGCTTCGGGAGAGAA	420
Db	361	CTTCCCGGGCGCGCATGATGGCGCGACAGGCGCGCGAGCGCGGCTTCGGGAGAGAA	420
QY	421	AGCGAGGGGGGAGGCGCGGAGCTCGCGGGACCTCGAGAGGGACCCCGCGCGAGAGAA	480
Db	421	AGCGAGGGGGGAGGCGCGGAGCTCGCGGGACCTCGAGAGGGACCCCGCGCGAGAGAA	480
QY	481	GCGGCGACCCCGGGGGCCACCGGGCTCGCGGGAGTCTCGGTGCTCTCCGAGATCCC	540
Db	481	GCGGCGACCCCGGGGGCCACCGGGCTCGCGGGAGTCTCGGTGCTCTCCGAGATCCC	540
QY	541	CTTCAGGGCCAAAGCGCTCCGAGAGCCGGTGCCTCGCGCTTGACGACACCTTGGCCCT	600
Db	541	CTTCAGGGCCAAAGCGCTCCGAGAGCCGGTGCCTCGCGCTTGACGACACCTTGGCCCT	600
QY	601	CGACCGCGTCTGCTTAAGACAGACGAGATTAACAGCGCGTACACCGGCAATTACCTG	660
Db	601	CGACCGCGTCTGCTTAAGACAGACGAGATTAACAGCGCGTACACCGGCAATTACCTG	660
QY	661	CGAGGCGCTTGGGGTCTACTACTTCCGCGTCCATGCGACCGCTACCGGCGCAGCTGCA	720
Db	661	CGAGGCGCTTGGGGTCTACTACTTCCGCGCTCCATGCGACCGCTACCGGCGCAGCTGCA	720
QY	721	GTTTGATCTGTGAAGAATGGGAAATCCATTGCGCTTCTTCCAGTTTTCGGGGGTG	780
Db	721	GTTTGATCTGTGAAGAATGGGAAATCCATTGCGCTTCTTCCAGTTTTCGGGGGTG	780
QY	781	GCCCCAAGCCAGCCTCGCTCGGGGGGGGCCATGTGTGAGGCTGAGACCTGAGACCAAGT	840
Db	781	GCCCCAAGCCAGCCTCGCTCGGGGGGGGCCATGTGTGAGGCTGAGACCTGAGACCAAGT	840
QY	841	GTTGGTGCAGTGGGGTGGGTGATCTACATTGCGCATTTATGCGACGATCAAGACAGACAG	900
Db	841	GTTGGTGCAGTGGGGTGGGTGATCTACATTGCGCATTTATGCGACGATCAAGACAGACAG	900
QY	901	CACCTTCTCCGGATTTCTGTGTACTCCGATGGCAGAGCTCCCGACGCTTGTGTTAGTG	960
Db	901	CACCTTCTCCGGATTTCTGTGTACTCCGATGGCAGAGCTCCCGACGCTTGTGTTAGTG	960
QY	961	CCCACTGGAAAGTAGATGATGCTCTCACTCTAGAAAGGAGGTGAGGCTGACAACA	1020
Db	961	CCCACTGGAAAGTAGATGATGCTCTCACTCTAGAAAGGAGGTGAGGCTGACAACA	1020
QY	1021	GCTCATCCAGAGGAGGCTGGCCCCCTGGAAATTGTGAATTACTAAGGAGGTGGGGTAGA	1080
Db	1021	GCTCATCCAGAGGAGGCTGGCCCCCTGGAAATTGTGAATTACTAAGGAGGTGGGGTAGA	1080
QY	1081	GCACCTCCCGCTCTGCTCTGTGGCAAGAAATGGAAAGTGGCTTCTGCGATCAGCTCTG	1140
Db	1081	GCACCTCCCGCTCTGCTCTGTGGCAAGAAATGGAAAGTGGCTTCTGCGATCAGCTCTG	1140
QY	1141	GCACCATGGGCACTGGCTGGATTTCTGGCCAAAGACAGAGAGTGTCTGTCTGGCAAA	1200
Db	1141	GCACCATGGGCACTGGCTGGATTTCTGGCCAAAGACAGAGAGTGTCTGTCTGGCAAA	1200
QY	1201	GTTTAAAGTCCCGCAGTTGCTGTGGTCCAGAGAGCCAGCGAGGGGTGCTCTTCCCTGATG	1260
Db	1201	GTTTAAAGTCCCGCAGTTGCTGTGGTCCAGAGAGCCAGCGAGGGGTGCTCTTCCCTGATG	1260
QY	1261	CTGCGCTTCTCTGATCTCTCCCAACCCCGCTCTGCTCTGGGCGCGGCGCTTCTCTAGA	1320
Db	1261	CTGCGCTTCTCTGATCTCTCCCAACCCCGCTCTGCTCTGGGCGCGGCGCTTCTCTAGA	1320

```

Oy      1321  GATCATCTCAATAAAGCTTAAGAACCCCTCATTAATAAAAAAAAAAAAAAAAAAAAAA 13777
Db      1321  GATCACTCAATAAACCTTAAGAACCCCTCATTAATAAAAAAAAAAAAAAAAAAAAAA 13777

RESULT 3
US-09-944-896-41
: Sequence 41, Application US/09944896
: Patent No. US20020168715A1
GENERAL INFORMATION:
: APPLICANT: Baker, Kevin
: APPLICANT: Botstein, David
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroff, Ellen
: APPLICANT: Gertlisen, Mary
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul
: APPLICANT: Grimaldi, Christopher
: APPLICANT: Gurney, Austin
: APPLICANT: Hillan, Kenneth
: APPLICANT: Kljavin, Ivar
: APPLICANT: Napier, Mary
: APPLICANT: Roy, Margaret
: APPLICANT: Tunas, Daniel
: APPLICANT: Wood, William
TITLE OF INVENTION: ACIDS AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P2348P1C1
CURRENT FILING DATE: US/09/944, 896
PRIORITY FILING DATE: 2001-08-31
PRIORITY APPLICATION NUMBER: 09/866, 028
PRIORITY FILING DATE: 2001-05-25
PRIORITY APPLICATION NUMBER: 60/069, 334
PRIORITY FILING DATE: December 11, 1997
PRIORITY APPLICATION NUMBER: 60/069335
PRIORITY FILING DATE: December 11, 1997
PRIORITY APPLICATION NUMBER: 60/069, 278
PRIORITY FILING DATE: December 11, 1997
PRIORITY APPLICATION NUMBER: 60/069, 425
PRIORITY FILING DATE: December 12, 1997
PRIORITY APPLICATION NUMBER: 60/069, 696
PRIORITY FILING DATE: December 16, 1997
PRIORITY APPLICATION NUMBER: 60/069, 694
PRIORITY FILING DATE: December 16, 1997
PRIORITY APPLICATION NUMBER: 60/069, 702
PRIORITY FILING DATE: December 16, 1997
PRIORITY APPLICATION NUMBER: 60/069, 870
PRIORITY FILING DATE: December 17, 1997
PRIORITY APPLICATION NUMBER: 60/069, 873
PRIORITY FILING DATE: December 17, 1997
PRIORITY APPLICATION NUMBER: 60/068, 017
PRIORITY FILING DATE: December 18, 1997
PRIORITY APPLICATION NUMBER: 60/070, 440
PRIORITY FILING DATE: January 5, 1998
PRIORITY APPLICATION NUMBER: 60/074, 086
PRIORITY FILING DATE: February 9, 1998
PRIORITY APPLICATION NUMBER: 60/074, 092
PRIORITY FILING DATE: February 9, 1998
PRIORITY APPLICATION NUMBER: 60/075, 945
PRIORITY FILING DATE: February 25, 1998
PRIORITY APPLICATION NUMBER: 60/111, 850
PRIORITY FILING DATE: December 16, 1998
PRIORITY APPLICATION NUMBER: 60/113, 286
PRIORITY FILING DATE: December 22, 1998
PRIORITY APPLICATION NUMBER: 60/146, 222
PRIORITY FILING DATE: July 28, 1999
PRIORITY APPLICATION NUMBER: PCT/US98/19930
PRIORITY FILING DATE: September 16, 1998
PRIORITY APPLICATION NUMBER: PCT/US98/25108
PRIORITY FILING DATE: December 1, 1998
PRIORITY APPLICATION NUMBER: 09/216, 021
PRIORITY FILING DATE: December 16, 1998

```

```

: PRIOR APPLICATION NUMBER: 09/218,517
: PRIOR FILING DATE: December 22, 1998
: PRIOR APPLICATION NUMBER: 09/254,311
: PRIOR FILING DATE: March 3, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/12252
: PRIOR FILING DATE: June 22, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/21090
: PRIOR FILING DATE: September 15, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/28409
: PRIOR FILING DATE: No. US20020168715A1ember 30, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/28313
: PRIOR FILING DATE: No. US20020168715A1ember 30, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/28301
: PRIOR FILING DATE: December 1, 1999
: PRIOR APPLICATION NUMBER: PCT/US99/30095
: PRIOR FILING DATE: December 16, 1999
: PRIOR APPLICATION NUMBER: PCT/US00/03565
: PRIOR FILING DATE: February 11, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/04414
: PRIOR FILING DATE: February 22, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/05841
: PRIOR FILING DATE: March 2, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/08439
: PRIOR FILING DATE: March 30, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/14042
: PRIOR FILING DATE: May 22, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/20710
: PRIOR FILING DATE: July 28, 2000
: PRIOR APPLICATION NUMBER: PCT/US00/32678
: PRIOR FILING DATE: December 1, 2000
: PRIOR APPLICATION NUMBER: PCT/US01/06520
: PRIOR FILING DATE: February 28, 2001
: NUMBER OF SEQ ID NOS: 120
: SEQ ID NO 41
: LENGTH: 1377
: TYPE: DNA
: ORGANISM: Homo Sapien
US-09-944-896-41

```

```

Query Match      100.0%  Score 1377:  DB 9:  Length 1377:
Best Local Similarity 100.0%  Pred. No. 0:
Matches 1377:  Conservative 0:  Mismatches 0:  Indels 0:  Gaps 0:

```

```

QY 1 GACTACTTCTTGTGAGTGTGGAGAGAGAAACCGAGCCGCGAGGAGCAACAGAC 60
DB 1 GACTACTTCTTGTGAGTGTGGAGAGAGAAACCGAGCCGCGAGGAGCAACAGAC 60
QY 1 TGGGGTGGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 1 TGGGGTGGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 121 CCAACTGAGAGGTCCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 121 CCAACTGAGAGGTCCTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
QY 181 GGGAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
DB 181 GGGAGTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 240
QY 241 CGTCTGCTGTCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 300
DB 241 CGTCTGCTGTCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCTGTGGGCT 300
QY 301 CAGCCCTGCTCCGCGGAGCAACCCGCGCTTCAGAGCAAGCGCGGAGCAATGAGCCAGG 360
DB 301 CAGCCCTGCTCCGCGGAGCAACCCGCGCTTCAGAGCAAGCGCGGAGCAATGAGCCAGG 360
QY 361 CTTGCGGCGGCGGAGTGGCGGAGAGCGCGGAGAGCGCGGCGGCGGCGGCGGAGAA 420
DB 361 CTTGCGGCGGCGGAGTGGCGGAGAGCGCGGAGAGCGCGGCGGCGGCGGCGGAGAA 420
QY 421 AGCGGAGGCGGAGGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGAGAA 480
DB 421 AGCGGAGGCGGAGGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGAGAA 480

```

```

DB 421 AGCGGAGGCGGAGGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGAGAA 480
QY 481 GCGGAGAGCGGCGGAGGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGAGAA 540
DB 481 GCGGAGAGCGGCGGAGGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGAGAA 540
QY 541 CTTGAGGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGAGAA 600
DB 541 CTTGAGGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGCGGAGAGAGAA 600
QY 601 CGAGCGGCGGCTGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
DB 601 CGAGCGGCGGCTGTGTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
QY 661 CCAAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
DB 661 CCAAGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
QY 721 GTTGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
DB 721 GTTGTATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
QY 781 GCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 840
DB 781 GCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAGGCGGAG 840
QY 841 GTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
DB 841 GTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
QY 901 CACCTTCTCGGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
DB 901 CACCTTCTCGGATTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
QY 961 CCCACTGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
DB 961 CCCACTGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1020
QY 1021 GGTTCATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 GGTTCATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 GACATCTCCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
DB 1081 GACATCTCCGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
QY 1141 GCAGCATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1141 GCAGCATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1201 GTGTAGTCCGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
DB 1201 GTGTAGTCCGCGGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260
QY 1261 CTCTGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
DB 1261 CTCTGCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320
QY 1321 GATCAGTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377
DB 1321 GATCAGTCAATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1377

```

```

RESULT 4
US-09-944-944-41
: Sequence 41, Application US/0944944
: Patent No. US20020173463A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin
: APPLICANT: Batstein, David
: APPLICANT: Eaton, Dan
: APPLICANT: Ferrara, Napoleone
: APPLICANT: Filvaroff, Ellen

```

```

1  APPLICANT: Geriltsen, Mary
2  APPLICANT: Goddard, Audrey
3  APPLICANT: Godowski, Paul
4  APPLICANT: Grimaldi, Christopher
5  APPLICANT: Gurney, Austin
6  APPLICANT: Hillan, Kenneth
7  APPLICANT: Hillan, Ivar
8  APPLICANT: Napier, Mary
9  APPLICANT: Roy, Margaret
10 APPLICANT: Tumas, Daniel
11 APPLICANT: Wood, William
12
13 TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
14 FILE OF INVENTION: ACIDS ENCODING THE SAME
15 FILE REFERENCE: P2548PICI
16
17 CURRENT APPLICATION NUMBER: US/09/944,944
18
19 CURRENT FILING DATE: 2001-09-26
20
21 PRIOR APPLICATION NUMBER: 09/866,028
22
23 PRIOR FILING DATE: 2001-05-25
24
25 PRIOR APPLICATION NUMBER: 60/067,411
26
27 PRIOR FILING DATE: December 3, 1997
28
29 PRIOR APPLICATION NUMBER: 60/069,334
30
31 PRIOR FILING DATE: December 11, 1997
32
33 PRIOR APPLICATION NUMBER: 60/063335
34
35 PRIOR FILING DATE: December 11, 1997
36
37 PRIOR APPLICATION NUMBER: 60/069,278
38
39 PRIOR FILING DATE: December 11, 1997
40
41 PRIOR APPLICATION NUMBER: 60/069,425
42
43 PRIOR FILING DATE: December 12, 1997
44
45 PRIOR APPLICATION NUMBER: 60/069,686
46
47 PRIOR FILING DATE: December 16, 1997
48
49 PRIOR APPLICATION NUMBER: 60/069,694
50
51 PRIOR FILING DATE: December 16, 1997
52
53 PRIOR APPLICATION NUMBER: 60/069,702
54
55 PRIOR FILING DATE: December 16, 1997
56
57 PRIOR APPLICATION NUMBER: 60/069,870
58
59 PRIOR FILING DATE: December 17, 1997
60
61 PRIOR APPLICATION NUMBER: 60/069,873
62
63 PRIOR FILING DATE: December 17, 1997
64
65 PRIOR APPLICATION NUMBER: 60/068,017
66
67 PRIOR FILING DATE: December 18, 1997
68
69 PRIOR APPLICATION NUMBER: 60/070,440
70
71 PRIOR FILING DATE: January 5, 1998
72
73 PRIOR APPLICATION NUMBER: 60/074,086
74
75 PRIOR FILING DATE: February 9, 1998
76
77 PRIOR APPLICATION NUMBER: 60/074,092
78
79 PRIOR FILING DATE: February 9, 1998
80
81 PRIOR APPLICATION NUMBER: 60/075,945
82
83 PRIOR FILING DATE: February 25, 1998
84
85 PRIOR APPLICATION NUMBER: 60/112,850
86
87 PRIOR FILING DATE: December 16, 1998
88
89 PRIOR APPLICATION NUMBER: 60/113,296
90
91 PRIOR FILING DATE: December 22, 1998
92
93 PRIOR APPLICATION NUMBER: 60/146,222
94
95 PRIOR FILING DATE: July 28, 1999
96
97 PRIOR APPLICATION NUMBER: PCT/US98/19330
98
99 PRIOR FILING DATE: September 16, 1998
100
101 PRIOR APPLICATION NUMBER: PCT/US98/25108
102
103 PRIOR FILING DATE: December 1, 1998
104
105 PRIOR APPLICATION NUMBER: 09/216,021
106
107 PRIOR FILING DATE: December 16, 1998
108
109 PRIOR APPLICATION NUMBER: 09/218,517
110
111 PRIOR FILING DATE: December 22, 1998
112
113 PRIOR APPLICATION NUMBER: 09/254,311
114
115 PRIOR FILING DATE: March 3, 1999
116
117 PRIOR APPLICATION NUMBER: PCT/US99/12252
118
119 PRIOR FILING DATE: June 22, 1999
120
121 PRIOR APPLICATION NUMBER: PCT/US99/21090
122
123 PRIOR FILING DATE: September 15, 1999
124
125 PRIOR APPLICATION NUMBER: PCT/US99/28409
126
127 PRIOR FILING DATE: NO. US200201134651 September 30, 1999
128
129 PRIOR APPLICATION NUMBER: PCT/US99/28313
130
131 PRIOR FILING DATE: NO. US200201134651 September 30, 1999
132
133 PRIOR APPLICATION NUMBER: PCT/US99/28301

```

```

? PRIOR FILING DATE: December1, 1999
? PRIOR APPLICATION NUMBER: PCT/US99/30095
? PRIOR FILING DATE: December 16, 1999
? PRIOR APPLICATION NUMBER: PCT/US00/03565
? PRIOR FILING DATE: February 11, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/04414
? PRIOR FILING DATE: February 22, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/05841
? PRIOR FILING DATE: March 2, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/08439
? PRIOR FILING DATE: March 30, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/14042
? PRIOR FILING DATE: May 22, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/20710
? PRIOR FILING DATE: July 28, 2000
? PRIOR APPLICATION NUMBER: PCT/US00/32878
? PRIOR FILING DATE: December 1, 2000
? PRIOR APPLICATION NUMBER: PCT/US01/06520
? PRIOR FILING DATE: February 28, 2001
? NUMBER OF SEQ ID NOS: 120
? SEQ ID NO 41
? LENGTH: 1377
? TYPE: DNA
? ORGANISM: Homo Sapien
US-09-944-944-41

Query Match      100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY    1   GATTAATTCTCTTGGAGCTCTGGGAGAGAAACCGAGCCGAGGAGCAGAACCAAGCAC 60
DB    1   GACTAATTCTCTTGGAGCTCTGGGAGAGAAACCGAGCCGAGGAGCAGAACCAAGCAC 60

QY    61  TGGGCTGACGCGCAGGCGCAGGGGGCGGCTGGCGGGGAGAAGCGGGGGCTGTGAGCACCA 120
DB    61  TGGCGGTAGCGCAGCGAGGCGCAGGGGGCGGCTGGCGGGGAGAAGCGGGGGCTGTGAGCACCA 120

QY    121 CCNACTGGAGGCTCCGGAAGTAGCAGACGCCCGGAAAGAGACCATCGGGAGCGCGGAGG 180
DB    121 CCNAAGTGAGGCTCCGGAAGTAGCAGACGCCCGGAAAGAGACCATCGGGAGCGCGGAGG 180

QY    181 GGGACTGCAGAGAGAACCCCGGGGCTCCGGGCTCCGGTCACACGCTATAGGCACACTCT 240
DB    181 GGGACTGCAGAGAGAACCCCGGGGCTCCGGGCTCCGGTCACACGCTATAGGCACACTCT 240

QY    241 CGTCTGCTGCTCTCTGGGCTTGGCGCGGCTTGGCCCCCATCTGAGACGACACMAATCCC 300
DB    241 CGTCTGCTGCTCTCTGGGCTTGGCGCGGCTTGGCCCCCATCTGAGACGACACMAATCCC 300

QY    301 CAGCCTTGCCCGGGGGACAACCCCGGCTTCCAGGCAACGCCCGGGGCACCATGCGACGCAGG 360
DB    301 CAGCCTTGCCCGGGGGACAACCCCGGCTTCCAGGCAACGCCCGGGGCACCATGCGACGCAGG 360

QY    361 CTTCGCGGGGCGGATTTGGCCCGCGAGAGCGCGCGAGACGCCCGGGGGGCTCCGGGAGAAA 420
DB    361 CTTCGCGGGGCGGATTTGGCCCGCGAGAGCGCGCGAGAGCGCGCGGGGCTCCGGGAGAAA 420

QY    421 AGGCGAGGGCGGGAGAGCCCGGCACTCTCGGAGACCTTCGAGGGAGACCCCGGGCGCGAGAGA 480
DB    421 AGGCGAGGGCGGGAGAGCCCGGCACTCTCGGAGACCTTCGAGGGAGACCCCGGGCGCGAGAGA 480

QY    481 GCGCGGAGCCCGGGGGGCCACACGGGGCGTGGCGGGGAGAGCTCGGGTGCCTCGCGCATCCGC 540
DB    481 GCGCGGAGCCCGGGGGGCCACACGGGGCGTGGCGGGGAGAGCTCGGGTGCCTCGCGCATCCGC 540

QY    541 CTTTAGCGCCCAAGGCTTCGAGAGCCCGGCTCGCTCGCGCTTCGAGACGACACTTTGGCCTT 600
DB    541 CTTTAGCGCCCAAGGCTTCGAGAGCCCGGCTCGCTCGCGCTTCGAGAGCAACCTTTGCCCTT 600

QY    601 CGACCGCGTCTGTTGTAAGCAGACAGGAGATTTAGAGAGCCGCTACACCGGCAAGTTACCTG 660
DB    601 CGACCGCGTCTGTTGTAAGCAGACAGGAGATTTAGAGAGCCGCTACACCGGCAAGTTACCTG 660

```

QY 661 CCAGTCCCTGGGCTACTACTGCGCCCTCCATGCCACCGTCTACCGGGCCAGCCCTCA 720
DB 661 CCAGTCCCTGGGCTACTACTGCGCCCTCCATGCCACCGTCTACCGGGCCAGCCCTCA 720
QY 721 GTTGTATCTGTGAAGATGGGGAATCCATTGCGCTTTCTTCCAGTTTTCGGGGGTG 780
DB 721 GTTGTATCTGTGAAGATGGGGAATCCATTGCGCTTTCTTCCAGTTTTCGGGGGTG 780
QY 781 GCCCAAGCCAGCTCTCTCGGGGGGGCCATGTGTAGAGCTGAGACCTGAGACCAAGT 840
DB 781 GCCCAAGCCAGCTCTCTCGGGGGGGCCATGTGTAGAGCTGAGACCTGAGACCAAGT 840
QY 841 GTGGGTGAGGTGGGTGGGTGAGTACATGAGCTATGCGACATCAAGACAGACAG 900
DB 841 GTGGGTGAGGTGGGTGGGTGAGTACATGAGCTATGCGACATCAAGACAGACAG 900
QY 901 CACCTTCTCCGATTTCTGTGTACTCCGACAGCTCCCGAGTCTTGTGTAGT 960
DB 901 CACCTTCTCCGATTTCTGTGTACTCCGACAGCTCCCGAGTCTTGTGTAGT 960
QY 961 CCCACTGGAAGTGAAGCTCATGCTCTCTCTAGAGAGAGGTGTAGAGCTGACAAACA 1020
DB 961 CCCACTGGAAGTGAAGCTCATGCTCTCTCTAGAGAGAGGTGTAGAGCTGACAAACA 1020
QY 1021 GGTATCCAGAGAGGCTGCGCCCTGGATATTTGTGATGAGGAGGTGGGTAGA 1080
DB 1021 GGTATCCAGAGAGGCTGCGCCCTGGATATTTGTGATGAGGAGGTGGGTAGA 1080
QY 1081 GCATCTCTCCGCTGCTGCTGCTGCAAGAGATGGAGAGAGGTGTGCTGCAATCAAGTCTG 1140
DB 1081 GCATCTCTCCGCTGCTGCTGCTGCAAGAGATGGAGAGAGGTGTGCTGCAATCAAGTCTG 1140
QY 1141 GCACATATGGGCGAGTGGCTGATTTCTGCCAAGACCAAGAGAGTGTGCTGCTGCAAA 1200
DB 1141 GCACATATGGGCGAGTGGCTGATTTCTGCCAAGACCAAGAGAGTGTGCTGCTGCAAA 1200
QY 1201 GTTGAATCCCGGAGTGTGCTGCTGCAAGAGAGAGGTGTGCTGCTGCTGCTGCTG 1260
DB 1201 GTTGAATCCCGGAGTGTGCTGCTGCAAGAGAGAGGTGTGCTGCTGCTGCTGCTG 1260
QY 1261 CTCTGCTTCTCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1261 CTCTGCTTCTCTGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
QY 1321 GATCAGTCAATTAACCTTAAGACCTCTATTAACCTCTATTAACCTCTATTAACCT 1377
DB 1321 GATCAGTCAATTAACCTTAAGACCTCTATTAACCTCTATTAACCTCTATTAACCT 1377

RESULT 5

US-09-944-907-41
Sequence 41, Application US/09944907
Publication No. US20020198147A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin
APPLICANT: Botstein, David
APPLICANT: Baton, Dan
APPLICANT: Ferrara, Napoleone
APPLICANT: Fillardoff, Ellen
APPLICANT: Geriltsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul
APPLICANT: Grimaud, Christopher
APPLICANT: Gurney, Austin
APPLICANT: Hillen, Kenneth
APPLICANT: Kijavlin, Ivar
APPLICANT: Napier, Mary
APPLICANT: Roy, Margaret
APPLICANT: Tumes, Daniel
APPLICANT: Wood, William
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME

FILE REFERENCE: P254BP1
CURRENT APPLICATION NUMBER: US/09/944, 907
CURRENT FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: 09/866, 028
PRIOR FILING DATE: 2001-05-25
NUMBER OF SEQ ID NOS: 120
SEQ ID NO 41
LENGTH: 1377
TYPE: DNA
ORGANISM: Homo Sapien
US-09-944-907-41
Query Match: 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GACTAGTCTCTTGTGAGTCTGGAGAGAGAAAGCCGAGCCGAGAGAGCAAGCAAGAC 60
DB 1 GACTAGTCTCTTGTGAGTCTGGAGAGAGAAAGCCGAGCCGAGAGAGCAAGCAAGAC 60
QY 61 TGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
DB 61 TGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 121 CCAACTGAGAGGTCGCGAGTACGAGCCGCGAAGAGAGGAGGAGGAGGAGGAGGAGG 180
DB 121 CCAACTGAGAGGTCGCGAGTACGAGCCGCGAAGAGAGGAGGAGGAGGAGGAGGAGG 180
QY 181 GGGAGTCCAG 240
DB 181 GGGAGTCCAG 240
QY 241 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 CGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 CAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 CAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 CTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
DB 361 CTTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 420
QY 421 AGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
DB 421 AGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
QY 481 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
DB 481 GCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
QY 541 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
DB 541 CTTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 600
QY 601 CGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
DB 601 CGACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
QY 661 CCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
DB 661 CCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
QY 721 GTTGTATCTGTGAAGATGGGGAATCCATTGCGCTTTCTTCCAGTTTTCGGGGGTG 780
DB 721 GTTGTATCTGTGAAGATGGGGAATCCATTGCGCTTTCTTCCAGTTTTCGGGGGTG 780
QY 781 GCCCAAGCCAGCTCTCTCGGGGGGGCCATGTGTAGAGCTGAGACCTGAGACCAAGT 840
DB 781 GCCCAAGCCAGCTCTCTCGGGGGGGCCATGTGTAGAGCTGAGACCTGAGACCAAGT 840
QY 841 GTGGGTGAGGTGGGTGGGTGAGTACATGAGCTATGCGACATCAAGACAGACAG 900

Query Match	100.08;	Score 1377;	DB 9;	Length 1377;
Best Local Similarity	100.08;	Pred. NO. 0;		

[illegible]

```
OY 1081 GCACTCTCCCTCTGCTGCTGCAAGAAAGTGGTCTGCTGCAAGTCTG 1140
Db 1081 GCACTCTCCCTCTGCTGCTGCAAGAAAGTGGTCTGCTGCAAGTCTG 1140
OY 1141 GCACATATGGGAGTGGTCTGATTTCTGCCAAGACAGAGAGTGTGCTGCTGCA 1200
Db 1141 GCACATATGGGAGTGGTCTGATTTCTGCCAAGACAGAGAGTGTGCTGCTGCA 1200
OY 1201 GTGTAATCTCCCGAGTGTGCTGCTGCAAGAGCCAGGAGGCTGCTCTTCTGCTC 1260
Db 1201 GTGTAATCTCCCGAGTGTGCTGCTGCAAGAGCCAGGAGGCTGCTCTTCTGCTC 1260
OY 1261 CTCTGCTCTCTGATCTCTCCCGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTG 1320
Db 1261 CTCTGCTCTCTGATCTCTCCCGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTG 1320
OY 1321 GATCCTCATTAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1377
Db 1321 GATCCTCATTAATTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1377

RESULT 7
US-10-028-072-361
; Sequence 361, Application US/10028072
; Publication No. US2003000431A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerltzen, Marcy E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang
; TITLE OF INVENTION:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/10/028, 072
; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/049911
; PRIOR FILING DATE: 1997-06-18
; PRIOR APPLICATION NUMBER: 60/056974
; PRIOR FILING DATE: 1997-08-26
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059115
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059117
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059122
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059184
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059263
; PRIOR FILING DATE: 1997-09-18
; PRIOR APPLICATION NUMBER: 60/059352
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059588
; PRIOR FILING DATE: 1997-09-19
; PRIOR APPLICATION NUMBER: 60/059836
; PRIOR FILING DATE: 1997-09-24
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062285

; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/062814
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/062816
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063045
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063082
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/063127
; PRIOR FILING DATE: 1997-10-24
; PRIOR APPLICATION NUMBER: 60/063327
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063329
; PRIOR FILING DATE: 1997-10-27
; PRIOR APPLICATION NUMBER: 60/063550
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063561
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/063704
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063733
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063735
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063738
; PRIOR FILING DATE: 1997-10-29
; PRIOR APPLICATION NUMBER: 60/063755
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064248
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/064809
; PRIOR FILING DATE: 1997-11-07
; PRIOR APPLICATION NUMBER: 60/065186
; PRIOR FILING DATE: 1997-11-12
; PRIOR APPLICATION NUMBER: 60/065846
; PRIOR FILING DATE: 1997-11-17
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/066453
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066511
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/066770
; PRIOR FILING DATE: 1997-11-24
; PRIOR APPLICATION NUMBER: 60/069212
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069278
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069334
; PRIOR FILING DATE: 1997-12-11
; PRIOR APPLICATION NUMBER: 60/069694
; PRIOR FILING DATE: 1997-12-16
; PRIOR APPLICATION NUMBER: 60/072320
; PRIOR FILING DATE: 1998-01-23
; PRIOR APPLICATION NUMBER: 60/073612
; PRIOR FILING DATE: 1998-02-04
; PRIOR APPLICATION NUMBER: 60/074086
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/074092
; PRIOR FILING DATE: 1998-02-09
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-02-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
```

PRIOR APPLICATION NUMBER: 60/080165
 PRIOR FILING DATE: 1998-03-31
 PRIOR APPLICATION NUMBER: 60/081203
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081229
 PRIOR FILING DATE: 1998-04-09
 PRIOR APPLICATION NUMBER: 60/081695
 PRIOR FILING DATE: 1998-04-14
 PRIOR APPLICATION NUMBER: 60/081817
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/081818
 PRIOR FILING DATE: 1998-04-15
 PRIOR APPLICATION NUMBER: 60/082999
 PRIOR FILING DATE: 1998-04-24
 PRIOR APPLICATION NUMBER: 60/083322
 PRIOR FILING DATE: 1998-04-28
 PRIOR APPLICATION NUMBER: 60/083545
 PRIOR FILING DATE: 1998-04-29
 PRIOR APPLICATION NUMBER: 60/084600
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084627
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/084637
 PRIOR FILING DATE: 1998-05-07
 PRIOR APPLICATION NUMBER: 60/085149
 PRIOR FILING DATE: 1998-05-12
 PRIOR APPLICATION NUMBER: 60/085323
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085338
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085339
 PRIOR FILING DATE: 1998-05-13
 PRIOR APPLICATION NUMBER: 60/085579
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085697
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/085704
 PRIOR FILING DATE: 1998-05-15
 PRIOR APPLICATION NUMBER: 60/086414
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/086430
 PRIOR FILING DATE: 1998-05-22
 PRIOR APPLICATION NUMBER: 60/087106
 PRIOR FILING DATE: 1998-05-28
 PRIOR APPLICATION NUMBER: 60/088026
 PRIOR FILING DATE: 1998-06-04
 PRIOR APPLICATION NUMBER: 60/088730
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088741
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088810
 PRIOR FILING DATE: 1998-06-10
 PRIOR APPLICATION NUMBER: 60/088858
 PRIOR FILING DATE: 1998-06-11
 PRIOR APPLICATION NUMBER: 60/089532
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089599
 PRIOR FILING DATE: 1998-06-17
 PRIOR APPLICATION NUMBER: 60/089907
 PRIOR FILING DATE: 1998-06-18
 PRIOR APPLICATION NUMBER: 60/089947
 PRIOR FILING DATE: 1998-06-19
 PRIOR APPLICATION NUMBER: 60/090349
 PRIOR FILING DATE: 1998-06-23
 PRIOR APPLICATION NUMBER: 60/090429
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090445
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090538
 PRIOR FILING DATE: 1998-06-24
 PRIOR APPLICATION NUMBER: 60/090863
 PRIOR FILING DATE: 1998-06-26
 PRIOR APPLICATION NUMBER: 60/091360

PRIOR FILING DATE: 1998-07-01
 PRIOR APPLICATION NUMBER: 60/091519
 PRIOR FILING DATE: 1998-07-02
 PRIOR APPLICATION NUMBER: 60/091982
 PRIOR FILING DATE: 1998-07-07

Query Match 100.0%; Score 1377; DB 9; Length 1377;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACTAGTCTCTTGTGATCTGGGAGGAGAAACCGGAGCCGCGAGAGGAGCGAAGAC 60
 1 GACTAGTCTCTTGTGATCTGGGAGGAGAAACCGGAGCCGCGAGAGGAGCGAAGAC 60
 61 TGGGATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
 61 TGGGATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
 121 CCAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
 121 CCAACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
 181 GGGACTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
 181 GGGACTGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
 241 GGTCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 241 GGTCTGCTGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
 301 CAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 301 CAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
 361 CTTGCGGCG 420
 361 CTTGCGGCG 420
 421 AGCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
 421 AGCGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480
 481 GCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 481 GCGGAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 540
 541 CTTGAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
 541 CTTGAGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
 601 CGAGCGGCTGCTGTGTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
 601 CGAGCGGCTGCTGTGTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
 661 CGAGGCTGCTGCTGTGTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 661 CGAGGCTGCTGCTGTGTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 720
 721 GTTGATCTGTGTAAGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
 721 GTTGATCTGTGTAAGATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 780
 781 GCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
 781 GCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
 841 GTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
 841 GTGGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 900
 901 CACCTTCTCGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
 901 CACCTTCTCGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960

[illegible]

```

RESULT 8
US-10-121-049-361
; Sequence 361, Application US/10121049
; Publication No. US20030022239A1
GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C17
; CURRENT APPLICATION NUMBER: US/10/121,049
; CURRENT FILING DATE: 2002-04-12
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-121-049-361

```

Query Match	100.0%	Score 1377:	DB 9:	Length 1377:
Best Local Similarity	100.0%	Prid. NO. 0:		
Matches 1377:	Conservative	Mismatches 0:	Indels 0:	Gaps 0:
1	GACTGTTCTCTTGAGTCTGGAGCGAAGAACGGCAGCGCAGCAGACGAAACCGAGAC	60		
1	GACTGTTCTCTTGAGTCTGGAGCGAAGAACGGCAGCGCAGCAGACGAAACCGAGAC	60		

QY	61	TGGGGTGACGGCAGAGGCGACAGGGGGCGCTTGCCCGGGGAAACGCGGGGGCTGTGAGCAACA	120
Db	61	TGGGGTGACGGCAGAGGCGACAGGGGGCGCTTGCCCGGGGAGAAACGCGGGGGCTGTGAGCAACA	120
QY	121	CCAACTGGAAGGGTCCGAGTAGTACGAGCGCCCCCGAAAGAGGCGCATTCGGGGAGACC	180
Db	121	CCAACTGGAAGGGTCCGAGTAGTACGAGCGCCCCCGAAAGAGGCGCATTCGGGGAGACC	180
QY	181	GGGACTGCGAAGAGACACCCCGGCGTCCGGGGCTCCCGGTGCCAGCGCTATGAGCGACACT	240
Db	181	GGGACTGCGAAGAGAGACCCCGGCGTCCGGGGCTCCCGGTGCCAGCGCTATGAGCGACACT	240
QY	241	CGTCTGCTGCTCTCTGGAGCCCTGGGGGGGGGGCTGCCCCCACTGAGAGCAACAAGATCCC	300
Db	241	CGTCTGCTGCTCTCTGGAGCCCTGGGGGGGGGGCTGCCCCCACTGAGAGCAACAAGATCCC	300
QY	301	CAGCCTCTGCCCGGGGACCCCGGCTTCCAGAGCAGCCCGGCGCACCATTGGAGCCAGAGG	360
Db	301	CAGCCTCTGCCCGGGGACCCCGGCTTCCAGAGCAGCCCGGCGCACCATTGGAGCCAGAGG	360
QY	361	CTTCCCGGGGCGCGATGGCCGCGACAGCGCGCGACGGCGCGCCCGGGGCTCCGGAGAGA	420
Db	361	CTTCCCGGGGCGCGATGGCCGCGACAGCGCGCGACGGCGCGCCCGGGGCTCCGGAGAGA	420
QY	421	AGCGAGAGGCGGAGAGCGCGGACTGCGCGGACCTCAGAGGAGACCCCGGGCGGAGAGAGA	480
Db	421	AGCGAGAGGCGGAGAGCGCGGACTGCGCGGACCTCAGAGGAGACCCCGGGCGGAGAGAGA	480
QY	481	GGGGGAGCCCGGGGGGCCACCGGGGCGTCCGGGGAGTCTCGGTCCGCGATCCGC	540
Db	481	GGGGGAGCCCGGGGGGCCACCGGGGCGTCCGGGGAGTCTCGGTCCGCGATCCGC	540
QY	541	CTTACGCGCCAAAGCGCTCCGAGAGCGGGTGCCTCGCGCTGTGAGAGACCCCTTCCCT	600
Db	541	CTTACGCGCCAAAGCGCTCCGAGAGCGGGTGCCTCGCGCTGTGAGAGACCCCTTCCCT	600
QY	601	CGACCGCGTGGTGGTGAACAGACAGAGGAGATTACAGACGCGCTACCGGCAAGTTTCACTG	660
Db	601	CGACCGCGTGGTGGTGGTGAACAGACAGAGGAGATTACAGACGCGCTACCGGCAAGTTTCACTG	660
QY	661	CCAGGTGCTTGGGGTCTACTACTTGGCGGTCCATCCACACCGCTACGGGGGACAGCTGCA	720
Db	661	CCAGGTGCTTGGGGTCTACTACTTGGCGGTCCATCCACACCGCTACGGGGGACAGCTGCA	720
QY	721	GTTTATCTGCTGGAAGAAATGCGGAATCCATTTCTTTCAGAGTTTTCGGGGGGTG	780
Db	721	GTTTATCTGCTGGAAGAAATGCGGAATCCATTTCTTTCAGAGTTTTCGGGGGGTG	780
QY	781	GCCCAAGCAGACCTTCGCTCTCGGGGGGGGGCCATGTGAGCGCTGAGAGCCAACT	840
Db	781	GCCCAAGCAGACCTTCGCTCTCGGGGGGGGGCCATGTGAGCGCTGAGAGCCAACT	840
QY	841	GTGGGTGAGGTGGGTGGTGTGACTACATTTGGCATCTATGCGAGCATCAAGACAGACAG	900
Db	841	GTGGGTGAGGTGGGTGGTGTGACTACATTTGGCATCTATGCGAGCATCAAGACAGACAG	900
QY	901	CACCTTCCCGGATTTTCTGTGTACTCCGATGGCACAGTCCCACTTTTCTTAACTG	960
Db	901	CACCTTCCCGGATTTTCTGTGTACTCCGATGGCACAGTCCCACTTTTCTTAACTG	960
QY	961	CCCACTGCAAAAGTAGCTCATGCTCTACACTAGAAAGAGGGTGTGAGGCTGACAAACA	1020
Db	961	CCCACTGCAAAAGTAGCTCATGCTCTACACTAGAAAGAGGGTGTGAGGCTGACAAACA	1020
QY	1021	GGTATCAGAGAGGGCTGGCGCCCGCTGGAAATTTGTGAAATGACTAGGAGAGTGGGTAGA	1080
Db	1021	GGTATCAGAGAGGGCTGGCGCCCGCTGGAAATTTGTGAAATGACTAGGAGAGTGGGTAGA	1080
QY	1081	GCACTCTCGTCTGTGCTGTGGCAAGAAATGGACAAGTGGCTGTGTGAGATCAGACTTG	1140
Db	1081	GCACTCTCGTCTGTGCTGTGGCAAGAAATGGACAAGTGGCTGTGTGAGATCAGACTTG	1140
QY	1141	GCAGCAATGGGGCGAGTGGTGATTTTGCCCAAGACACAGAGATGTGCTGTGGCA	1200

Db 1141 GAGAGATGGGGCAGTGGCTGATTTCTGCCCCAAGACAGAGAGTGTGCTGTCGGACA 1200
QY 1201 GTGTAAGTCCCGCCAGTGTGTCGTCAGAGCCACAGCTGGGGTGTCTCTCTGCTC 1260
Db 1201 GTGTAAGTCCCGCCAGTGTGTCGTCAGAGCCACAGCTGGGGTGTCTCTCTGCTC 1260
QY 1261 CTCTGCTCTCTGATTCCTCCCGACCCCTCTCTGCTGCTGGGGCCGCTTTTCTCAGA 1320
Db 1261 CTCTGCTCTCTGATTCCTCCCGACCCCTCTCTGCTGCTGGGGCCGCTTTTCTCAGA 1320
QY 1321 GATCAGCTCAATTAACCTTAAGAACCTCTATAAAAAAAAAAAAAAAAAAAAAA 1377
Db 1321 GATCAGCTCAATTAACCTTAAGAACCTCTATAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 9
US-10-123-904-361
Sequence 361, Application US/10123904
Publication No. US20030022328A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Collin K.
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P330R1C34
CURRENT APPLICATION NUMBER: US/10/123,904
CURRENT FILING DATE: 2002-04-16
Prior Application removed - See File Wrapper or Palm
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 361
LENGTH: 1377
TYPE: DNA
ORGANISM: Homo Sapien
US-10-123-904-361

Query Match: 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 241 GGTCTGCTGCTCTCTGGGCTTGGGGCCGCTGCCCCCACTGAGCAACAAGATCCC 300
QY 301 CAGCCTTGCCCGGGGACACCCCGGCTTCCAGGACAGCCGGGACCATGAGCCAGAG 360
Db 301 CAGCCTTGCCCGGGGACACCCCGGCTTCCAGGACAGCCGGGACCATGAGCCAGAG 360
QY 361 CTTCGCGGGCGGGGATGAGCGGCGGACAGCGCGGCGGCGGCGGCGGCGGCGGAGAA 420
Db 361 CTTCGCGGGCGGGGATGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGAA 420
QY 421 AGGCGAGGGGCGGAGAGCGCGGAGCTGCGGAGCTCGAGGGGAGCCCGGGCGGAGAGA 480
Db 421 AGGCGAGGGGCGGAGAGCGCGGAGCTGCGGAGCTCGAGGGGAGCCCGGGCGGAGAGA 480
QY 481 GCGGGAGCCCGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
Db 481 GCGGGAGCCCGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 541 CTTCAGGCGCCAGCGCTCGGAGAGCGGGGCTCGGCGGCTCGGCGGCTGAGAGCACCTTGCCTT 600
Db 541 CTTCAGGCGCCAGCGCTCGGAGAGCGGGGCTCGGCGGCTCGGCGGCTGAGAGCACCTTGCCTT 600
QY 601 CGACCGCGCTGCTGTAAGAGAGAGGACATTACAGCGCGCTACCGGCAAGTTCACCTG 660
Db 601 CGACCGCGCTGCTGTAAGAGAGAGGACATTACAGCGCGCTACCGGCAAGTTCACCTG 660
QY 661 CCAGGTCCTGGGGTACTACTTGGCGCTCATGCGACCGTCAACCGGGCCAGCCTGCA 720
Db 661 CCAGGTCCTGGGGTACTACTTGGCGCTCATGCGACCGTCAACCGGGCCAGCCTGCA 720
QY 721 GTTGTATCTGTAAGATGCGGAATTCATGCTCTTCTTCCAGTTTTCGGGGGGG 780
Db 721 GTTGTATCTGTAAGATGCGGAATTCATGCTCTTCTTCCAGTTTTCGGGGGGG 780
QY 781 GCCCAAGCCAGCCTGCTCTCGGGGGGCGCATGTGAGCTGAGACCTGAGACCAAGT 840
Db 781 GCCCAAGCCAGCCTGCTCTCGGGGGGCGCATGTGAGCTGAGACCTGAGACCAAGT 840
QY 841 GTGGGTCAGTGGGGTGGTGGTACTACTATGGCATGATGACACATCAAGACAGACAG 900
Db 841 GTGGGTCAGTGGGGTGGTGGTACTACTATGGCATGATGACACATCAAGACAGACAG 900
QY 901 CACCTTTCGCGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
Db 901 CACCTTTCGCGGATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
QY 961 CCCACTGCAAGTGAAGTGCATGCTCTCACTCTCTAAGAGAGGCTGAGAGCTGACACCA 1020
Db 961 CCCACTGCAAGTGAAGTGCATGCTCTCACTCTCTAAGAGAGGCTGAGAGCTGACACCA 1020
QY 1021 GGTCAATCCAGAGGCTGCGCCCGCTGGAATATTGTAATGACTAGGAGGTCGGGTAGA 1080
Db 1021 GGTCAATCCAGAGGCTGCGCCCGCTGGAATATTGTAATGACTAGGAGGTCGGGTAGA 1080
QY 1081 GCACCTCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
Db 1081 GCACCTCTCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 GCAGCATGGGGCAGTGGCTGGAATTTCTGCGCAAGCAAGAGAGTGTCTGCTGCGCA 1200
Db 1141 GCAGCATGGGGCAGTGGCTGGAATTTCTGCGCAAGCAAGAGAGTGTCTGCTGCGCA 1200
QY 1201 GTGTAAGTCCCGCCAGTGTGTCGTCAGAGCCACAGCTGGGTGCTCTTCTGCTGTC 1260
Db 1201 GTGTAAGTCCCGCCAGTGTGTCGTCAGAGCCACAGCTGGGTGCTCTTCTGCTGTC 1260
QY 1261 CTCTGCTCTCTGATTCCTCCCGACCCCTCTCTGCTGCTGGGGCCGCTTTTCTCAGA 1320
Db 1261 CTCTGCTCTCTGATTCCTCCCGACCCCTCTCTGCTGCTGGGGCCGCTTTTCTCAGA 1320
QY 1321 GATCAGCTCAATTAACCTTAAGAACCTCTATAAAAAAAAAAAAAAAAAAAAAA 1377
Db 1321 GATCAGCTCAATTAACCTTAAGAACCTCTATAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 10

```

US-10-140-470-361
; Sequence 361, Application US/10140470
; Publication No. US20030022331A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Collin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P33081C160
; CURRENT APPLICATION NUMBER: US/10/140,470
; PRIOR APPLICATION removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 361
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-140-470-361

```

Query Match 100.0%; Score 1377; DB 9; Length 1377;

Best Local Similarity 100.0%; Pred. No. 0; Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 GACTAGTCTCTTGGAGTCTGGAGAGGAAAGCCGAGCGGAGCGAAGCAACGAGAC 60
DB 1 GACTAGTCTCTTGGAGTCTGGAGAGGAAAGCCGAGCGGAGCGAAGCAACGAGAC 60
QY 61 TGGGGTGACGCGAGGAGGCGGCTGCGCGGGGAGAGCGGGGGGCTGAGACCA 120
DB 61 TGGGGTGACGCGAGGAGGAGGCGGCTGCGCGGGGAGAGCGGGGGGCTGAGACCA 120
QY 121 CCAACTGAGAGGTCGCGAGTACGAGCGCCCGAAGAGGCAATCGGGAGCGGAGAGG 180
DB 121 CCAACTGAGAGGTCGCGAGTACGAGCGCCCGAAGAGGCAATCGGGAGCGGAGAGG 180
QY 181 GGGAGTGCAGAGAGAGCCCGGCGCTCGGGGCTCCGGGTGCAAGGCTATAGGCACTCT 240
DB 181 GGGAGTGCAGAGAGAGCCCGGCGCTCGGGGCTCCGGGTGCAAGGCTATAGGCACTCT 240
QY 241 CGTCTGCTGCTCTTGGGCGCTGGGCGCGCTGCGCCCACTGAGACAAAGATGCC 300
DB 241 CGTCTGCTGCTCTTGGGCGCTGGGCGCGCTGCGCCCACTGAGACAAAGATGCC 300
QY 301 CAGCCTCTGCGCGGCGACCCCGGCTTCAGAGCAAGCGGCGCACATGCAAGCCAGGG 360
DB 301 CAGCCTCTGCGCGGCGACCCCGGCTTCAGAGCAAGCGGCGCACATGCAAGCCAGGG 360
QY 361 CTTCGCGGCGCGAGTGCAGCGAGCGCGAGCGCGCGCGGCGCTCCGCGAGAGA 420
DB 361 CTTCGCGGCGCGAGTGCAGCGAGCGCGAGCGCGCGCGGCGCTCCGCGAGAGA 420
QY 421 AGCGGAGGCGGAGGCGCGGAGTGCAGCGAGCGCGAGCGCGCGCGGCGCTCCGCGAGAGA 480
DB 421 AGCGGAGGCGGAGGCGCGGAGTGCAGCGAGCGCGAGCGCGCGCGGCGCTCCGCGAGAGA 480

```

```

QY 481 GCGCGAGCCCGCGGCGCCCAACCGGCGCTGCGGGGAGTGTCTCGGTCTCCGAGTCCG 540
DB 481 GCGCGAGCCCGCGGCGCCCAACCGGCGCTGCGGGGAGTGTCTCGGTCTCCGAGTCCG 540
QY 541 CTTCAGCGCAAGGCTCCGAGAGCGGCGGTGCTCCGCGCTGAGACACCTTGCCTT 600
DB 541 CTTCAGCGCAAGGCTCCGAGAGCGGCGGTGCTCCGCGCTGAGACACCTTGCCTT 600
QY 601 CGACCGGCTGCTGTGAGAGAGAGGACATTACAGCGCGCTACCGGCAAGTTCACTG 660
DB 601 CGACCGGCTGCTGTGAGAGAGGAGACATTACAGCGCGCTACCGGCAAGTTCACTG 660
QY 661 CAGAGTGCCTGGGGTCTACTACTTCCCGGCAATGCGACCGGCTACCGGCGCAAGT 720
DB 661 CAGAGTGCCTGGGGTCTACTACTTCCCGGCAATGCGACCGGCTACCGGCGCAAGT 720
QY 721 GTTTGATCTGTGTAAGAAATGCGAATCATTCCTCTTCTTCAGATTTCGGGGGTG 780
DB 721 GTTTGATCTGTGTAAGAAATGCGAATCATTCCTCTTCTTCAGATTTCGGGGGTG 780
QY 781 GCCCAAGCCAGCCTGCTCTCGGGGGGCGCATGTGAGAGCTGAGACCTGAGACCAAGT 840
DB 781 GCCCAAGCCAGCCTGCTCTCGGGGGGCGCATGTGAGAGCTGAGACCTGAGACCAAGT 840
QY 841 GTGGGTGAGAGTGGGTGTGTGACTACATTTGAGCATCTATGACACATCAAGACAGAG 900
DB 841 GTGGGTGAGAGTGGGTGTGTGACTACATTTGAGCATCTATGACACATCAAGACAGAG 900
QY 901 CACCTTCGCGGATTTTGTGTGTACTCCGACATGCGACAGCTCCCGAGTCTTCTTACTG 960
DB 901 CACCTTCGCGGATTTTGTGTGTACTCCGACATGCGACAGCTCCCGAGTCTTCTTACTG 960
QY 961 CCCACTGCAAACTGAGCTCATCTCACTCTAGAGAGAGGTGAGGTGAGCAACCA 1020
DB 961 CCCACTGCAAACTGAGCTCATCTCACTCTAGAGAGAGGTGAGGTGAGCAACCA 1020
QY 1021 GGTTCATCCAGAGGCGGTGCGCCCTGGAATATTGTGAATGAGAGGTGGGTGAGA 1080
DB 1021 GGTTCATCCAGAGGCGGTGCGCCCTGGAATATTGTGAATGAGAGGTGGGTGAGA 1080
QY 1081 GCACCTCCGCTCTCTCTGCGCAAGAAATGGGAACACTGGCTCTCCGATCAGAGTCTG 1140
DB 1081 GCACCTCCGCTCTCTCTGCGCAAGAAATGGGAACACTGGCTCTCCGATCAGAGTCTG 1140
QY 1141 GCAGCATGGGCGAGTGTGATTTTCCCAAGACAGAGAGTGTGTCTGTGCGAA 1200
DB 1141 GCAGCATGGGCGAGTGTGATTTTCCCAAGACAGAGAGTGTGTCTGTGCGAA 1200
QY 1201 GTGTAAATCCCGCAAGTTGCTGTCCAGAGAGCCAGGCTGGGTGCTCTCTGCTGTC 1260
DB 1201 GTGTAAATCCCGCAAGTTGCTGTCCAGAGAGCCAGGCTGGGTGCTCTCTGCTGTC 1260
QY 1261 CTCTGCTCTCTGTGATCTCTCCCAACCCCTCTCTCTGCGGCGCGCTTTTCTAGA 1320
DB 1261 CTCTGCTCTCTGTGATCTCTCCCAACCCCTCTCTCTGCGGCGCGCTTTTCTAGA 1320
QY 1321 GATCACTCAATAAAGCTTAAGAACCTCTATAAAAAAAAAAAAAAAAAAAAAA 1377
DB 1321 GATCACTCAATAAAGCTTAAGAACCTCTATAAAAAAAAAAAAAAAAAAAAAA 1377

```

RESULT 11

```

US-10-175-746-361
; Sequence 361, Application US/10175746
; Publication No. US2003002270A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.

```

APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C353
CURRENT APPLICATION NUMBER: US/10/175,746
PRIORITY FILING DATE: 2002-06-19
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 361
LENGTH: 1377
TYPE: DNA
ORGANISM: Homo Sapien
US-10-175-746-361

Query Match 100.0% Score 1377: DB 9: Length 1377:

Best Local Similarity 100.0% Pred. No. 0: Mismatches 0: Indels 0: Gaps 0:

Matches 1377: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

QY 1 GACTAGTCTCTTGGAGTCTGGAGAGAGAAAGCCGAGCCGAGAGAGCAACCAAGAGC 60
DB 1 GACTAGTCTCTTGGAGTCTGGAGAGAGAAAGCCGAGCCGAGAGAGCAACCAAGAGC 60
QY 61 TGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
DB 61 TGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 120
QY 121 CCAACTGAGAGGCTCCGAGATAGCAGAGCCGCCGAGAGGAGGAGGAGGAGGAGGAGGAGG 180
DB 121 CCAACTGAGAGGCTCCGAGATAGCAGAGCCGCCGAGAGGAGGAGGAGGAGGAGGAGGAGG 180
QY 181 GGGAGTGGAG 240
DB 181 GGGAGTGGAG 240
QY 241 GGTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
DB 241 GGTCTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300
QY 301 CAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
DB 301 CAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 360
QY 361 CTTGGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
DB 361 CTTGGCGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 420
QY 421 AGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
DB 421 AGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 480
QY 481 GGGCGGAGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
DB 481 GGGCGGAGCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 540
QY 541 CTTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
DB 541 CTTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 600
QY 601 CGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
DB 601 CGACCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 660
QY 661 CAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
DB 661 CAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720

DB 661 CAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 720
QY 721 GTTGTATCTGTGAAGAAATGGGAAATCCATTCCTCTTCTCCAGATTTTTCGGGGGTG 780
DB 721 GTTGTATCTGTGAAGAAATGGGAAATCCATTCCTCTTCTCCAGATTTTTCGGGGGTG 780
QY 781 GGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
DB 781 GGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 840
QY 841 GTGGGTGAGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 900
DB 841 GTGGGTGAGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 900
QY 901 CACCTTCTCCGAAATTTCTGTGTACTCCGAGTGGCAGAGTCCCACTCTTCTTCTTCTTCT 960
DB 901 CACCTTCTCCGAAATTTCTGTGTACTCCGAGTGGCAGAGTCCCACTCTTCTTCTTCTTCT 960
QY 961 CCACTGCAAGTGAAGTCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
DB 961 CCACTGCAAGTGAAGTCAATGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 1021 GGTATCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 GGTATCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 GCACTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
DB 1081 GCACTCTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
QY 1141 GCACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
DB 1141 GCACATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
QY 1201 GTGTAACTCCCAAGT 1260
DB 1201 GTGTAACTCCCAAGT 1260
QY 1261 CTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
DB 1261 CTGCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1320
QY 1321 GATCACTCAATTAACCTTAAGAACCTTCATTAACCTTCATTAACCTTCATTAACCTTCAT 1377
DB 1321 GATCACTCAATTAACCTTAAGAACCTTCATTAACCTTCATTAACCTTCATTAACCTTCAT 1377

RESULT 12
US-10-176-918-361
Sequence 361, Application US/10176918
Publication No. US20030027275A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Oiang
APPLICANT: Gerltsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330R1C382
CURRENT APPLICATION NUMBER: US/10/176,918


```
Db 1 GACTACTTCTCTGAGACTCTGGAGAGAAAGGAGCGGAGGAGGACGGAACCAAGAGC 60
QY 61 TGGGGTGGAGCGAGGCGAGGCGGCGCTGGCCGGGAGAAAGCGCGGGCTGTAGACACCA 120
Db 61 TGGGGTGGAGCGAGGCGAGGCGGCGCTGGCCGGGAGAAAGCGCGGGCTGTAGACACCA 120
QY 121 CCAACTGGAGGGGCGGAGTACGAGACGCCGCCGGAAGAGAGCCATGGGGAGCGGAGGG 180
Db 121 CCAACTGGAGGGGCGGAGTACGAGACGCCGCCGGAAGAGAGCCATGGGGAGCGGAGGG 180
QY 181 GGGAGTGGAGAGAGACCCGGCGCTCCGGGCTCCGGGTCCAGCGGTATGAGGCCACTCT 240
Db 181 GGGAGTGGAGAGAGACCCGGCGCTCCGGGCTCCGGGTCCAGCGGTATGAGGCCACTCT 240
QY 241 GGTCTGCTGCTCTGCTGGGCGCTGGGCGCGGCTGGCCCGCACTGAGAGCAACAGATCCC 300
Db 241 GGTCTGCTGCTCTGCTGGGCGCTGGGCGCGGCTGGCCCGCACTGAGAGCAACAGATCCC 300
QY 301 CAGCGCTGGCCCGGGGGGACCCCGGCTTCAGAGCACCGCGGCGCAATGGCAGCGCAGG 360
Db 301 CAGCGCTGGCCCGGGGGGACCCCGGCTTCAGAGCACCGCGGCGCAATGGCAGCGCAGG 360
QY 361 CTTGCGCGGCGCGGATGGCGCGGAGCGGCGGCGCGCGCGCGCGCGCTCCGGAGAGAA 420
Db 361 CTTGCGCGGCGCGGATGGCGCGGAGCGGCGGCGCGCGCGCGCGCGCTCCGGAGAGAA 420
QY 421 AGGCGAGGCGGCGGAGCGCGGCGGAGCTCCGGGACCTCGAGGGAGACCCGGCGCGGAGAGA 480
Db 421 AGGCGAGGCGGCGGAGCGCGGCGGAGCTCCGGGACCTCGAGGGAGACCCGGCGCGGAGAGA 480
QY 481 GCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGAGA 480
Db 481 GCGCGGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGAGA 480
QY 541 CTTGAGCGCCCAAGCGCTCCGAGAGCGCGGCGGCTCCCGCGCTGACAGCAGCCCTTGCCCTT 600
Db 541 CTTGAGCGCCCAAGCGCTCCGAGAGCGCGGCGGCTCCCGCGCTGACAGCAGCCCTTGCCCTT 600
QY 601 CGACCGCGCTGCTGTGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
Db 601 CGACCGCGCTGCTGTGAAGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 660
QY 661 CCAGGCGCTGGGCTCTACTACTTCCGCGTCAATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
Db 661 CCAGGCGCTGGGCTCTACTACTTCCGCGTCAATGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 720
QY 721 GTTGGATCTGGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
Db 721 GTTGGATCTGGTGAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 780
QY 781 GCGCAAGCGAGCGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGT 840
Db 781 GCGCAAGCGAGCGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGT 840
QY 841 GTGGGTCAGAGTGGGTCGAGTCTGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
Db 841 GTGGGTCAGAGTGGGTCGAGTCTGAGTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900
QY 901 CACCTTCTCCGAGATTTCTGTACTCCGATGCGAGCTCCCAAGTCTTCTCTTCTAGT 960
Db 901 CACCTTCTCCGAGATTTCTGTACTCCGATGCGAGCTCCCAAGTCTTCTCTTCTAGT 960
QY 961 CCCACTGCAAGAGTACGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
Db 961 CCCACTGCAAGAGTACGATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1020
QY 1021 GGTCTATTCAGAGAGGCGTGGCGCGGCTGGAGATATTGTAATGATAGGAGGAGGAGGAGGAGG 1080
Db 1021 GGTCTATTCAGAGAGGCGTGGCGCGGCTGGAGATATTGTAATGATAGGAGGAGGAGGAGGAGG 1080
QY 1081 GCACTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
Db 1081 GCACTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
```

```
Db 1081 GCACTCTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1140
QY 1141 GCAGCATGGGCGATGGCTGTGATTTCTTGCCCAAGACAGAGGAGTGTGTGTGCGCA 1200
Db 1141 GCAGCATGGGCGATGGCTGTGATTTCTTGCCCAAGACAGAGGAGTGTGTGTGCGCA 1200
QY 1201 GTGTAAGTCCCGCAAGTGTCTGTGTCCAGAGGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
Db 1201 GTGTAAGTCCCGCAAGTGTCTGTGTCCAGAGGCCACAGGAGGAGGAGGAGGAGGAGGAGGAGG 1260
QY 1261 CTCTGCTTCTCTGATCCCTCCCAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
Db 1261 CTCTGCTTCTCTGATCCCTCCCAAGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1320
QY 1321 GATCACTCAATAACCTTAAGACCCCTATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377
Db 1321 GATCACTCAATAACCTTAAGACCCCTATATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1377

RESULT 14
US-10-137-865-361
: Sequence 361, Application US/10137865
: Publication No. US2003003215A1
: GENERAL INFORMATION:
: APPLICANT: Baker, Kevin P.
: APPLICANT: Beresini, Maureen
: APPLICANT: Desnoyers, Luc
: APPLICANT: Flivaroff, Ellen
: APPLICANT: Gao, Wei-Qiang
: APPLICANT: Geriltsen, Mary E.
: APPLICANT: Goddard, Audrey
: APPLICANT: Godowski, Paul J.
: APPLICANT: Gurney, Austin L.
: APPLICANT: Sherwood, Steven
: APPLICANT: Smith, Victoria
: APPLICANT: Stewart, Timothy A.
: APPLICANT: Tumas, Daniel
: APPLICANT: Watanabe, Colin K
: APPLICANT: Wood, William
: APPLICANT: Zhang, Zemin
: TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
: FILE REFERENCE: ACIDS ENCODING THE SAME
: CURRENT APPLICATION NUMBER: US/10/137,865
: CURRENT FILING DATE: 2002-05-03
: Prior Application removed - See Palm or File Wrapper
: NUMBER OF SEQ ID NOS: 550
: SEQ ID NO 361
: LENGTH: 1377
: TYPE: DNA
: ORGANISM: Homo Sapien
US-10-137-865-361

Query Match 100.0%; Score 1377; DB 9; Length 1377;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GACTACTTCTCTGAGACTCTGGAGAGAAAGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
Db 1 GACTACTTCTCTGAGACTCTGGAGAGAAAGGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 60
QY 61 TGGGGTGGAGCGAGGCGAGGCGGCGCTGGCCGGGAGAAAGCGCGGGCTGTAGACACCA 120
Db 61 TGGGGTGGAGCGAGGCGAGGCGGCGCTGGCCGGGAGAAAGCGCGGGCTGTAGACACCA 120
QY 121 CCAACTGGAGGGTCCGAGTACGAGACGCCGCCGGAAGAGAGCCATGGGGAGCGGAGGG 180
Db 121 CCAACTGGAGGGTCCGAGTACGAGACGCCGCCGGAAGAGAGCCATGGGGAGCGGAGGG 180
QY 181 GGGAGTGGAGAGAGACCCGGCGCTCCGGGCTCCGGGTCCAGCGGTATGAGGCCACTCT 240
Db 181 GGGAGTGGAGAGAGACCCGGCGCTCCGGGCTCCGGGTCCAGCGGTATGAGGCCACTCT 240
```

OY	241	GTCTCTGCTGCTCTCTGGGCTTGCGGCCGGCGTCCGCCACCTGAGCACAACAAGATCCC	300
Db	241	GTCCTCTGCTGCTCTCTGGGCTTGCGGCCGGCGTCCGCCACCTGAGCACAACAAGATCCC	300
OY	301	CAGCCTCTGCCCCGGGGGACACCCGGGCTTTCAGGAGACCCGGGCGACCATGGAAGCCAGG	360
Db	301	CAGCCTCTGCCCCGGGGGACACCCGGGCTTTCAGGAGACCCGGGCGACCATGGAAGCCAGG	360
OY	361	CTTGCCGGGGCCCGCATGAGCCCGCAGACGGCGCGCAGAGCGCCCGGGGGCTCCGAGAGAA	420
Db	361	CTTGCCGGGGCCCGCATGAGCCCGCAGACGGCGCGCAGAGCGCCCGGGGGCTCCGAGAGAA	420
OY	421	AGGGAGGGGGCGGAGGGCGGGGACTGCGGGGAACTCCAGAGGGACCCGGGGCCCGAGAGAA	480
Db	421	AGGGAGGGGGCGGAGGGCGGGGAACTGCGGGGAACTCCAGAGGGACCCGGGGCCCGAGAGAA	480
OY	481	GGCGGGACCCGGGGGGCCCAACCGGGGCTGCGGGGAGTGTGTGAGTGTCTCCGAGATCCGC	540
Db	481	GGCGGGACCCGGGGGGCCCAACCGGGGCTGCGGGGAGTGTGTGAGTGTCTCCGAGATCCGC	540
OY	541	CTTCAGGGGGCAAGCGCTCCGAGAGCCGGGGTGCCTCCGCGTGTGAGCACCCTTGCCCTT	600
Db	541	CTTCAGGGGGCAAGCGCTCCGAGAGCCGGGGTGCCTCCGCGTGTGAGCACCCTTGCCCTT	600
OY	601	CGACCCGCTGCTGTGTGAGAGAGAGGACATTACAGACGCGTACCGCGCAACTTACCTG	660
Db	601	CGACCCGCTGCTGTGTGAGAGAGAGGACATTACAGACGCGTACCGCGCAACTTACCTG	660
OY	661	CCAGGGGCGCTGGGGGTCTACTACTTGCGCGCTCCATGCGCACCGCTACCGGGCCAGCTGCA	720
Db	661	CCAGGGGCGCTGGGGGTCTACTACTTGCGCGCTCCATGCGCACCGCTACCGGGCCAGCTGCA	720
OY	721	GTTTGATCTGTGTGAAGAAATGGGAAATCCATTCGCTTCTCCAGTTTTCGGGGGGTG	780
Db	721	GTTTGATCTGTGTGAAGAAATGGGAAATCCATTCGCTTCTCCAGTTTTCGGGGGGTG	780
OY	781	GGCCAAAGCCAGCCTGCGCTCTGGGGGGGGCCCATGTGTGAGGGCTGAGGACCAAGT	840
Db	781	GGCCAAAGCCAGCCTGCGCTCTGGGGGGGGCCCATGTGTGAGGGCTGAGGACCAAGT	840
OY	841	GTGGGTGCAGGTGGGTGTGGGTGATCTACATTTGGCATTCATTCGACGACATCAAGACAGAG	900
Db	841	GTGGGTGCAGGTGGGTGTGGGTGATCTACATTTGGCATTCATTCGACGACATCAAGACAGAG	900
OY	901	CACCTTCTCCGGATTTCGTGTACTCCGATGGCAGACGTCGCCAGCTCTTGCTTGTGTG	960
Db	901	CACCTTCTCCGGATTTCGTGTACTCCGATGGCAGACGTCGCCAGCTCTTGCTTGTGTG	960
OY	961	CCCACTGTGAAGTAGGCTCATGCTCTACTCTCTGAAAGAGAGGTGTGAGGCTGACAACA	1020
Db	961	CCCACTGTGAAGTAGGCTCATGCTCTACTCTCTGAAAGAGAGGTGTGAGGCTGACAACA	1020
OY	1021	GCTCATCCAGAGAGGCTGGCCCCCTCGAATATGTGAATGACATGAGGAGGTGGGGTAG	1080
Db	1021	GCTCATCCAGAGAGGCTGGCCCCCTCGAATATGTGAATGACATGAGGAGGTGGGGTAG	1080
OY	1081	GCACCTCCGCTCGCTGCTGTGGGCAAGAAATGGGAACGTGGCTGTGCGATCAGGCTGTG	1140
Db	1081	GCACCTCCGCTCGCTGCTGTGGGCAAGAAATGGGAACGTGGCTGTGCGATCAGGCTGTG	1140
OY	1141	GCACCATGGGGACATGGCTGATTTCTGCCCCAAGACACAGAGAGTGTGCTGTGGCA	1200
Db	1141	GCACCATGGGGACATGGCTGATTTCTGCCCCAAGACACAGAGAGTGTGCTGTGGCA	1200
OY	1201	GTGTAATTCGCCAAGTGTCTGTGTGTGAGAGCCACAGGTGGGGTGTCTCTTCCGTGTC	1260
Db	1201	GTGTAATTCGCCAAGTGTCTGTGTGTGAGAGCCACAGGTGGGGTGTCTCTTCCGTGTC	1260
OY	1261	CTGTGCTTCTCTGTGATCTTCCACCCCTCTGCTGCTGAGGGCGGGCCCTTTTCTCAGA	1320
Db	1261	CTGTGCTTCTCTGTGATCTTCCACCCCTCTGCTGCTGAGGGCGGGCCCTTTTCTCAGA	1320

QY	1321	GATCAGCTCAATTAACCTTAAGAACCCCTCATATAAAAAAAAAAAAAAAAAAAAAA	1377
Db	1321	GATCAGCTCAATTAACCTTAAGAACCCCTCATATAAAAAAAAAAAAAAAAAAAAAA	1377
		US-10-140-474-361	
		RESULT 15	
		US-10-140-474-361	
		Sequence 361: Application US/10140474	
		Publication No. US20030032156A1	
		GENERAL INFORMATION:	
		APPLICANT: Baker, Kevin P.	
		APPLICANT: Beresini, Maureen	
		APPLICANT: Desforger, Laura	
		APPLICANT: Desnoyers, Luc	
		APPLICANT: Flivaatoff, Ellen	
		APPLICANT: Gao, Wei-Qiang	
		APPLICANT: Gerritsen, Mary E.	
		APPLICANT: Goddard, Audrey	
		APPLICANT: Godowski, Paul J.	
		APPLICANT: Gunney, Austin L.	
		APPLICANT: Sherwood, Steven	
		APPLICANT: Smith, Victoria	
		APPLICANT: Stewart, Timothy A.	
		APPLICANT: Tumas, Daniel	
		APPLICANT: Watanabe, Collin K	
		APPLICANT: Wood, William	
		APPLICANT: Zhang, Zemin	
		TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC	
		FILE REFERENCE: P3330R1C162	
		CURRENT APPLICATION NUMBER: US/10/140,474	
		CURRENT FILING DATE: 2002-05-06	
		Prior Application removed - See Palm or File Wrapper	
		NUMBER OF SEQ ID NOS: 550	
		SEQ ID NO 361	
		LENGTH: 1377	
		TYPE: DNA	
		ORGANISM: Homo Sapien	
		US-10-140-474-361	
		Query Match	
		Best Local Similarity 100.0%; Score 1377; DB 9; Length 1377;	
		Matches 1377; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GACTGCTCTCTTGGAGTCTGGGAGGAGAAACCGAGCCGGCAGCAGCGAACCCAGAC	60
Db	1	GACTGCTCTCTTGGAGTCTGGGAGGAGAAACCGAGCCGGCAGCAGCGAACCCAGAC	60
QY	61	TGGGGTGACGCGAGGCGCAAGGGGGCCCTGGCCGGGGAGAAACCGGGGGCTGTGACACCA	120
Db	61	TGGGGTGACGCGAGGCGCAAGGGGGCCCTGGCCGGGGAGAAACCGGGGGCTGTGACACCA	120
QY	121	CCAACTGAGGGGTCCGAGTAGACGAGCGCCCGGAAGAGAGGCATCTGGGGAGCCGGGAGGG	180
Db	121	CCAACTGAGGGGTCCGAGTAGACGAGCGCCCGGAAGAGAGGCATCTGGGGAGCCGGGAGGG	180
QY	181	GGGACTGGAGAGAGACCCCGGGCGTCGGGGCTCCCGGTCCAGGCGCTTATGAGCCACTCTCT	240
Db	181	GGGACTGGAGAGAGACCCCGGGCGTCGGGGCTCCCGGTCCAGGCGCTTATGAGCCACTCTCT	240
QY	241	CGTCTCTGCTCTCTCTGGGCGCTGGGGCGCGCTCCGCCCACTGGAGCAGCAACAAGATCCC	300
Db	241	CGTCTCTGCTCTCTCTGGGCGCTGGGGCGCGCTCCGCCCACTGGAGCAGCAACAAGATCCC	300
QY	301	CAGCCTCTGGCCCGGGGAGCCCGGGCTTTCAGAGCAGCCGGGGCAGCATGCGACCCAGGG	360
Db	301	CAGCCTCTGGCCCGGGGAGCCCGGGCTTTCAGAGCAGCCGGGGCAGCATGCGACCCAGGG	360
QY	361	CTTTCGCGGGCGCGGATGCGCGCAGACGGCCGCGCCCGGGGGCTCTCGGGAGACAA	420
Db	361	CTTTCGCGGGCGCGGATGCGCGCAGACGGCCGCGCCCGGGGGCTCTCGGGAGACAA	420
QY	421	AGGCGAGGGCGGAGGCGCGGACTTCCGCGGAGACTTCAGAGGGAGACCCCGGGCGCGGAGAGA	480

DB 421 AGGCGAGGGCGGAGCGGCGGAGCTCGCGGAGCTGAGAGGAGACCCCGGCGGAGAGAG 480
QY 481 GGGGGGAGCGCGGG 540
DB 481 GGGGGGAGCGCGGG 540
QY 541 CTTGAGCGCGCAAGGCGCTCGAGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 600
DB 541 CTTGAGCGCGCAAGGCGCTCGAGAGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 600
QY 601 CGACCGCGCTGCTGAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
DB 601 CGACCGCGCTGCTGAGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 660
QY 661 CCAGGTCGCTGGGGTCTACTCTCGCGGTCGATGCGGCGGTCGACGCGGCGGCGGCGGCG 720
DB 661 CCAGGTCGCTGGGGTCTACTCTCGCGGTCGATGCGGCGGTCGACGCGGCGGCGGCGGCG 720
QY 721 GTTTGATCTGGTGAAGAATGCGCAATTCATTCCTCTCTCCAGTTTTCGGGGGGGG 780
DB 721 GTTTGATCTGGTGAAGAATGCGCAATTCATTCCTCTCTCCAGTTTTCGGGGGGGG 780
QY 781 GCGCAAGCGCGCTCGGCTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 840
DB 781 GCGCAAGCGCGCTCGGCTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 840
QY 841 GTGGGTGCAAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 900
DB 841 GTGGGTGCAAGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 900
QY 901 CACCTTCTCGGATTTCTGTGCTACTCCGACAGCTCCCGCAGTCTTTGCTTAGTG 960
DB 901 CACCTTCTCGGATTTCTGTGCTACTCCGACAGCTCCCGCAGTCTTTGCTTAGTG 960
QY 961 CCCACTGCAAAAGTACCTCATGCTCAGCTCCAGAAAGAGGAGGAGGAGGAGGAGGAGGAG 1020
DB 961 CCCACTGCAAAAGTACCTCATGCTCAGCTCCAGAAAGAGGAGGAGGAGGAGGAGGAGGAG 1020
QY 1021 GGTGATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
DB 1021 GGTGATCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
QY 1081 GCACCTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
DB 1081 GCACCTCCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1140
QY 1141 GCAGCATGGGGGAGTGGCTGATTTCTGCGCAAGACAGAGAGTGGTGGTGGCGCA 1200
DB 1141 GCAGCATGGGGGAGTGGCTGATTTCTGCGCAAGACAGAGAGTGGTGGTGGCGCA 1200
QY 1201 GTTGAAGTCCCGCAGTTGCTGTGTCAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
DB 1201 GTTGAAGTCCCGCAGTTGCTGTGTCAGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1260
QY 1261 CTTCTGCTTCTGATTCCTCCGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1320
DB 1261 CTTCTGCTTCTGATTCCTCCGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 1320
QY 1321 GATCACTCAATAAAGCTAAGAACCTCATATAAAAAAAAAAAAAAAAAAAAAA 1377
DB 1321 GATCACTCAATAAAGCTAAGAACCTCATATAAAAAAAAAAAAAAAAAAAAAA 1377